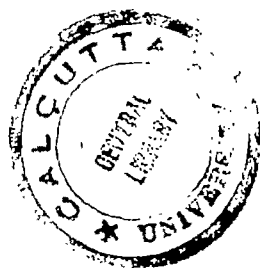


C 012 H03469-15-2001292

American Sociological Review

15



Night As Frontier Murray Melbin	3
Diffusion of Collective Violence Brian L. Pitcher, Robert L. Hamblin and Jerry L. L. Miller	23
Women's Intergenerational Occupational Mobility Rachel A. Rosenfeld	36
Curriculum Tracking and Educational Stratification Karl L. Alexander, Martha Cook and Edward L. McDill	47
Marx's Theory of the Falling Rate of Profit Richard P. Appelbaum	67
Residential Segregation by Occupation and Race Albert A. Simkus	81
Race and Involvement in Crimes Michael J. Hindelang	93
Comments Edward J. Walsh, James R. Lincoln and Jon Olson, Robert M. Hauser, Joel H. Levine	110

Notice to Contributors

The maximum length of an ASR paper is typically ten (10) printed pages or thirty (30) typed manuscript pages including space for tables, figures, and references. Due to space limitations, we must request contributors to conform to this norm as closely as possible.

To permit anonymity in the review of manuscripts, keep identifying material out of the manuscript. Attach a cover page giving authorship, institutional affiliation and acknowledgments, and provide only the title as identification on the manuscript and abstract.

ASA Multiple Submissions Policy: Submission of a manuscript to a professional journal clearly implies commitment to publish in that journal. The competition for journal space requires a great deal of time and effort on the part of editorial readers whose main compensation for this service is the opportunity to read papers prior to publication and the gratification associated with discharge of professional obligation. For these reasons, the American Sociological Association regards submission of a manuscript to a professional journal while that paper is under review by another journal as unacceptable.

SUBMISSION AND PREPARATION OF MANUSCRIPTS

1. Submit four (4) copies and retain the original for your files. Copies may be Xerox, mimeograph or multilith, but not carbons.
2. Enclose a stamped self-addressed postcard for acknowledgment of manuscript receipt. *Manuscripts will not be returned.* Please do not send your original copy.
3. All copy must be typed, doublespaced (including indented material, footnotes and references) on 8½ by 11 inch white opaque paper. Lines must not exceed six (6) inches. Margins must be a minimum of one inch.
4. Include four (4) copies of an abstract of no more than 150 words.
5. Type each table on a separate page. Insert a location note at the appropriate place in the text, e.g., "Table 2 about here." Send copies, retain originals.
6. Figures must be drawn in India ink on white paper. Send copies, retain originals.
7. Clarify all symbols with notes in the margins of the manuscript. Circle these and all other explanatory notes not intended for printing.
8. Footnotes should not be placed in the body of the text. Type them (doublespaced) and attach them as a separate appendix to the text. Number them consecutively throughout the text. Footnotes are to be kept to a minimum and used only for substantive observations. Source citations are made within the text rather than in footnotes.
9. Acknowledgments, credits and grant numbers are placed on the title page with an asterisk.

REFERENCE FORMAT

- In the text:** All source references are to be identified at the appropriate point in the text by the last name of the author, year of publication and pagination where needed. Identify subsequent citations of the same source in the same way as the first, not using "ibid.," "op.cit.," or "loc. cit." Examples follow:
 1. If author's name is in the text, follow it with year in parentheses. ["... Duncan (1959) ..."]
 2. If author's name is not in the text, insert, in parentheses, the last name and year, separated by a comma. ["... (cf. Gouldner, 1963) ..."]
 3. Pagination (without "p." or "pp.") follows year of publication after a colon. ["... Kuhn (1970:71)."]
 4. Give both last names for dual authors; for more than two use "et al." in the text. When two authors have the same last name, use identifying initials in the text. For institutional authorship, supply minimum identification from the beginning of the complete citation. ["... (U.S. Bureau of the Census, 1963:117) ..."]
 5. Separate a series of references with semicolons and enclose them within a single pair of parentheses. ["... (Burgess, 1968; Marwell et al., 1971; Cohen, 1962) ..."]
- In the appendix:** List all source citations by author, and within author by year of publication, in an appendix titled "References." The reference appendix must be complete and include all references in the text. The use of "et al." is not acceptable in the appendix; list the names of all authors. (See A. 4. for text format.)
If there is more than one reference to the same author and year, distinguish them by the letters, a, b, etc. added to the year. ["... Levy (1965a:331) ..."] Give the publisher's name in as brief a form as is fully intelligible. For example, John A. Wiley and Sons should be "Wiley." If the cited material is unpublished, use "forthcoming" with name of journal or publisher; otherwise use "unpublished."
Use no underlining, italics or abbreviations.
Examples follow:

- 1. Books:**

Jud, Gerald J., Edgar W. Mills, Jr. and Genevieve Walters Burch
1970 Ex-Pastors. Philadelphia: Pilgrim Press.

U.S. Bureau of the Census
1960 Characteristics of Population. Volume 1. Washington, D.C.: U.S. Government Printing Office.

Bernard, Claude
[1865] An Introduction to the Study of Experimental Medicine. Tr. Henry Copley Greene. New York: Dover.
- 2. Periodicals:**

Conger, Rand
Forth- "The effects of positive feedback on direction and amount of verbalization in a social setting." Pacific Sociological Review.

Merton, Robert K.
1963a "The ambivalence of scientists." Bulletin of The Johns Hopkins Hospital 112:77-97.

1963b "Resistance to the systematic study of multiple discoveries in science." European Journal of Sociology 4:237-82.
- 3. Collections:**

Davie, M.
1938 "The pattern of urban growth." Pp. 133-61 in G. Murdock (ed.), Studies in the Science of Society. New Haven: Yale University Press.

See recent issues for further examples.

Revised 1976 ASR

P 1892

AMERICAN SOCIOLOGICAL REVIEW

Official Journal of the American Sociological Association

FEBRUARY, 1978

VOLUME 43, NUMBER 1

Vol. 43

ARTICLES

Night As Frontier	Murray Melbin	3
The Diffusion of Collective Violence	Brian L. Pitcher, Robert L. Hamblin and Jerry L. L. Miller	23
Women's Intergenerational Occupational Mobility	Rachel A. Rosenfeld	36
Curriculum Tracking and Educational Stratification: Some Further Evidence	Karl L. Alexander, Martha Cook and Edward L. McDill	47
Marx's Theory of the Falling Rate of Profit: Towards a Dialectical Analysis of Structural Social Change	Richard P. Appelbaum	67
Residential Segregation by Occupation and Race in Ten Urbanized Areas, 1950-1970	Albert A. Simkus	81
Race and Involvement in Common Law Personal Crimes	Michael J. Hindelang	93

COMMENTS

Comment on Jenkins and Perrow (ASR April, 1977)		
On the Interaction between a Movement and Its Environment	Edward J. Walsh	110
Comment on Grimes, Bonjean, Lyon and Lineberry (ASR August, 1976)		
The Ecological Approach and Community Leadership	James R. Lincoln and Jon Olson	112
Comment on Butler (ASR October, 1976)		
On Inequality and the Military	Robert M. Hauser	115
Comment on McCann (ASR February, 1977)		
Comparing Models of Mobility	Joel H. Levine	118

Editor: RITA J. SIMÓN

Deputy Editors: CLARK McPHAIL AND ROSS M. STOLZENBERG

Associate Editors: MICHAEL T. AIKEN, ROBERT P. ALTHAUSER, PHILLIP F. BONACICH, ROBERT COLE, JUDY C. CORDER-BOLZ, FRED DAVIS, BEVERLY DUNCAN, JOHN FREEMAN, ERVING GOFFMAN, ROSABETH KANTER, ALAN C. KERCKHOFF, BARBARA LASLETT, PETER MANNING, WILLIAM M. MASON, MURRAY MELBIN, ANTHONY OBERSCHALL, WHITNEY POPE, SAMUEL PRESTON, DAVID SNYDER, BARBARA SOBIESZEK, AAGE SØRENSEN, GERALD SUTTLES, DON H. ZIMMERMAN

Copy Editor: MARY S. MANDER

Managing Editor: LINNA McDADE

Executive Officer: RUSSELL R. DYNES

SCOPE AND MISSION OF AMERICAN SOCIOLOGICAL REVIEW: *The American Sociological Review* publishes work of interest to the discipline in general, new theoretical developments, results of research that advances our understanding of the most fundamental social processes, important methodological innovations. Like other publications of the Association, emphasis is given to exceptional quality. Unlike the more specialized journals of the Association, *The American Sociological Review's* primary focus is on whatever has the most general bearing on the knowledge of society.

Concerning manuscripts, address: Editor, *American Sociological Review*, Department of Sociology, University of Illinois, Urbana, Illinois 61801. (217) 333-6515.

The American Sociological Review is published at 49 Sheridan Avenue, Albany, New York, bimonthly in February, April, June, August, October, and December.

Applications for permission to quote from this journal should be addressed to the Executive Office.

Copyright © 1978 American Sociological Association.

ISSN 0003-1224

Annual membership dues of the Association: Member and International Member, \$30-50; Student Member, \$15; Associate, \$20; Student Associate, \$10; International Associate, \$12.

Application for membership and payment of dues should be made to the Executive Office. Subscription rate for members, \$10; nonmembers, \$15; institutions and libraries, \$30; nonmember students, \$10. Single issue, \$4.

New subscriptions and renewals will be entered on a calendar year basis only.

Change of address: Six weeks advance notice to the Executive Office and old address as well as new are necessary for change of subscriber's address.

Concerning advertising, changes of address, and subscriptions, address: Executive Office, American Sociological Association, 1722 N Street, N.W., Washington, D.C. 20036.

Claims for undelivered copies must be made within the month following the regular month of publication. The publishers will supply the missing copies when losses have been sustained in transit and when the reserve stock will permit.

Second class postage paid at Albany, N. Y.

Members of the Council

Officers of the Association for the year 1978

AMOS H. HAWLEY, *President*

North Carolina, Chapel Hill

ALICE S. ROSSI, *Vice President*

University of Massachusetts, Amherst

HUBERT BLALOCK, *President-Elect*

University of Washington

CHARLES Y. GLOCK, *Vice President-Elect*

University of California, Berkeley

JAMES F. SHORT, JR., *Secretary*

Washington State University

RUSSELL R. DYNES, *Executive Officer*

Ohio State University

J. MILTON YINGER, *Past President*

Oberlin College

Elected-at-Large

ELISE BOULDING

University of Colorado

ERNEST Q. CAMPBELL

Vanderbilt University

HERBERT L. COSTNER

National Science Foundation

RICHARD J. HILL

University of Oregon

JOAN HUBER

University of Illinois

LEWIS M. KILLIAN

University of Massachusetts

DAVID MECHANIC

University of Wisconsin

JOAN MOORE

University of Wisconsin

PAMELA A. ROBY

University of Cal., Santa Cruz

IMMANUEL WALLERSTEIN

SUNY, Binghamton

WILLIAM F. WHYTE

Cornell University

MAURICE ZEITLIN

University of Cal., Los Angeles

University of Wisconsin

(Articles in the REVIEW are indexed in the Social Sciences Index, Public Affairs Information Service, Psychological Abstracts, Sociological Abstracts, Ayer's Guide, University Microfilms, Abstracts for Social Workers, International Political Science Abstracts, and United States Political Science Documents.)

PROCESSING FEE

A processing fee of \$10 is required for each paper submitted; such fees will be waived for student members of ASA. This reflects a policy of the ASA Council and Committee on Publications affecting all ASA journals. It is a reluctant response to the rapidly accelerating costs of manuscript reviewing. A check or money order, made payable to the American Sociological Association should accompany each submission. The fee must be paid in order for the review process to begin.

ITEMS

■ We want to thank MORRIS ZELDITCH, PHYLIS CAIRNS, and HARRIET FOSTER for their help in making the transition from Stanford to Illinois as smooth and uncomplicated as it was. Having Phyllis Cairns on the premises for a week saved us many hours of second guessing about procedural matters and introduced the new editorial staff to the business of running a journal under the best of conditions.

■ The new editorial and office staff consists of LINNA McDADE, office manager, and MARY MANDER, copy editor. CLYDENE MORGAN serves as assistant to both of them.

■ Two of my colleagues have graciously agreed to

serve as deputy editors: CLARK McPHAIL and ROSS M. STOLZENBERG. McPhail works primarily in the area of social interaction, collective behavior, and methodology of systematic observation. He received his Ph.D. from Michigan State University and was on the faculty at Beloit College and at the University of South Carolina before coming to the University of Illinois. Stolzenberg works primarily in the area of occupational structure and labor force participation. He received his doctorate from the University of Michigan and was on the faculty at Harvard and Johns Hopkins University before coming to Illinois.

■ The Publications Committee has appointed the fol-

(Continued on page 122)

AMERICAN SOCIOLOGICAL REVIEW

Official Journal of the American Sociological Association

APRIL, 1978

VOLUME 43, NUMBER 2

ARTICLES

- Equality, Success, and Social Justice in England and the United States Robert V. Robinson and Wendell Bell 125
- Informed Consent: Consequences for Response Rate and Response Quality in Social Surveys Eleanor Singer 144
- Determinants of Juvenile Court Dispositions: Ascriptive and Achieved Factors in Two Metropolitan Courts Lawrence E. Cohen and James R. Kluegel 162
- Formal and Substantive Voluntarism in the Work of Talcott Parsons: A Theoretical and Ideological Reinterpretation Jeffrey C. Alexander 177
- Community Structure and Industrial Conflict: An Analysis of Strike Activity in SMSAs James R. Lincoln 199
- Status Generalization: A Review and Some New Data Murray Webster, Jr. and James E. Driskell, Jr. 220
- Sex Differences in Children's Friendships Donna Eder and Maureen T. Haflinan 237
- Suburban Social Status: Persistence or Evolution? Avery M. Guest 251
- A Third Interpretation for the Generating Process of the Negative Binomial Distribution ... William W. Eaton and Andrée Fortin 264

COMMENTS

- Comment on Hirschi and Hindelang (ASR August, 1977)
The Meaning of IQ Delinquency Ronald L. Simons 268
- Comment on Robinson (ASR August, 1976)
Internal Political Organization and the World Economy of Income Inequality Steven Stack 271
- Comment on Silberman (ASR June, 1976)
Deterrence and Social Control Harold G. Grasmick and Steven D. McLaughlin 272
- Comment on Harris (ASR February, 1977)
Attribution, Goals, and Deviance Joseph Harry and Mary C. Sengstock 278

Editor: RITA J. SIMON

Deputy Editors: CLARK MCPHAIL AND ROSS M. STOLZENBERG

Associate Editors: MICHAEL T. AIKEN, ROBERT P. ALTHAUSER, PHILLIP F. BONACICH, ROBERT COLE, JUDY C. CORDER-BOLZ, FRED DAVIS, BEVERLY DUNCAN, JOHN FREEMAN, ERVING GOFFMAN, ROSABETH KANTER, ALAN C. KERCKHOFF, BARBARA LASLETT, PETER MANNING, WILLIAM M. MASON, MURRAY MELBIN, ANTHONY OBERSCHALL, WHITNEY POPE, SAMUEL PRESTON, DAVID SNYDER, BARBARA SOBIESZEK, AAGE SØRENSEN, GERALD SUTTLES, DON H. ZIMMERMAN

Copy Editor: MARY S. MANDER

Editorial Assistant: CLYDENE MORGAN

Managing Editor: LINNA McDADE

Executive Officer: RUSSELL R. DYNES

COPE AND MISSION OF AMERICAN SOCIOLOGICAL REVIEW: The *American Sociological Review* publishes work of interest to the discipline in general, new theoretical developments, results of research that advances our understanding of the most fundamental social processes, and important methodological innovations. Like other publications of the Association, emphasis is given to exceptional quality. Unlike the more specialized journals of the Association, the *American Sociological Review's* primary focus is on what has the most general bearing on the knowledge of society.

Concerning manuscripts, address: Editor, *American Sociological Review*, Department of Sociology, University of Illinois, Urbana, Illinois 61801. (217) 333-6515.

The *American Sociological Review* is published at 49 Sheridan Avenue, Albany, New York, bimonthly in February, April, June, August, October, and December.

Applications for permission to quote from this journal should be addressed to the Executive Office.

Copyright © 1978 American Sociological Association.

ISSN 0003-1224



Annual membership dues of the Association: Member and International Member, \$30-50; Student Member, \$15; Associate, \$20; Student Associate, \$10; International Associate, \$12.

Application for membership and payment of dues should be made to the Executive Office. Subscription rate for members, \$10; nonmembers, \$15; institutions and libraries, \$30; nonmember students, \$10. Single issue, \$4.

New subscriptions and renewals will be entered on a calendar year basis only.

Change of address: Six weeks advance notice to the Executive Office and old address as well as new are necessary for change of subscriber's address.

Concerning advertising, changes of address, and subscriptions, address: Executive Office, American Sociological Association, 1722 N Street, N.W., Washington, D.C. 20036.

Claims for undelivered copies must be made within the month following the regular month of publication. The publishers will supply the missing copies when losses have been sustained in transit and when the reserve stock will permit.

Second class postage paid at Washington, D.C. and additional mailing offices.

Members of the Council

Officers of the Association for the year 1978

AMOS H. HAWLEY, *President*

North Carolina, Chapel Hill

ALICE S. ROSSI, *Vice President*

University of Massachusetts, Amherst

HUBERT BLALOCK, *President-Elect*

University of Washington

CHARLES Y. GLOCK, *Vice President-Elect*

University of California, Berkeley

JAMES F. SHORT, JR., *Secretary*

Washington State University

RUSSELL R. DYNES, *Executive Officer*

Ohio State University

J. MILTON YINGER, *Past President*

Oberlin College

Elected-at-Large

ELISE BOULDING

University of Colorado

ERNEST Q. CAMPBELL

Vanderbilt University

HERBERT L. COSTNER

National Science Foundation

RICHARD J. HILL

University of Oregon

JOAN HUBER

University of Illinois

LEWIS M. KILLIAN

University of Massachusetts

DAVID MECHANIC

University of Wisconsin

JOAN MOORE

University of Wisconsin

PAMELA A. ROBY

University of Cal., Santa Cruz

IMMANUEL WALLERSTEIN

SUNY, Binghamton

WILLIAM F. WHYTE

Cornell University

MAURICE ZEITLIN

University of Cal., Los Angeles

University of Wisconsin

(Articles in the REVIEW are indexed in the Social Sciences Index, Public Affairs Information Service, Psychological Abstracts, Sociological Abstracts, Ayer's Guide, University Microfilms, Abstracts for Social Workers, International Political Science Abstracts, and United States Political Science Documents.)

PROCESSING FEE

A processing fee of \$10 is required for each paper submitted; such fees will be waived for student members of ASA. This reflects a policy of the ASA Council and Committee on Publications affecting all ASA journals. It is a reluctant response to the rapidly accelerating costs of manuscript reviewing. A check or money order, made payable to the American Sociological Association should accompany each submission. The fee must be paid in order for the review process to begin.

ITEMS

■ ROBERT V. ROBINSON (Equality, Success and Social Justice) is a Ph.D. candidate in the Department of Sociology at Yale University and Research Associate for the Center for Policy Research, New York City. He has completed (with Wendell Bell) a study of cultural and political nationalism in Jamaica, and (with Jonathan Kelley) an empirical test of class theories of Marx and Dahrendorf on American and British data. He is working on a comparative analysis of the intergenerational transmission of capital, authority, and status. WENDELL BELL is Professor of Sociology at Yale University and Director of the Comparative Sociology Training Program. He

is working (with Robert V. Robinson and others) a restudy of Jamaican leaders and on a theoretical synthesis of inequality and social justice. He coedited (with Walter E. Freeman) *Ethnicity and Nation-Building: Comparative, International, and Historical Perspectives* (Sage, 1974).

■ ELEANOR SINGER (Informed Consent) Senior Research Associate in the Graduate School of Journalism, Columbia University, and Editor of *Public Opinion Quarterly*. She is continuing research on survey research methods.

(Continued on page 28)

AMERICAN SOCIOLOGICAL REVIEW

Official Journal of the American Sociological Association

JUNE, 1978

VOLUME 43, NUMBER 3

ARTICLES

- The Causes and Cost of Racial Exclusion from Job Authority James R. Kluegel 285
Class Power and State Policy: The Case of Large Business Corporations, Labor Unions and
Governmental Redistribution in the American States
..... Alexander Hicks, Roger Friedland, and Edwin Johnson 302
Determinants and Behavioral Consequences of Psychological Modernity: Empirical Evidence
from Costa Rica Michael Armer and Larry Isaac 316
Behaviorism on Verstehen and Erklären John D. Baldwin and Janice I. Baldwin 335
Explaining the Low Use of Health Services by the Poor: Costs, Attitudes, or Delivery Systems?
..... Diana B. Dutton 348
Social Factors in Psychiatric Outcome: Toward the Resolution of Interpretive Controversies
..... R. Jay Turner and John W. Gartrell 368
The Sociogenesis of Psychological Disorder: Reexamining the Causal Issues with Longitudinal
Data Blair Wheaton 383
Organizational Change in the American Foreign Service, 1925-1965: The Utility of Cohort
Analysis Theodore L. Reed 404
Style as Social Process Barbara Rosenblum 422

COMMENT

- Comment on Hope (ASR June, 1975)
Facets and Flaws of Hope's Diamond Model James S. House 439

Editor: RITA J. SIMON

Deputy Editors: CLARK McPHAIL AND ROSS M. STOLZENBERG

Associate Editors: MICHAEL T. AIKEN, ROBERT P. ALTHAUSER, PHILLIP F.
BONACICH, ROBERT COLE, JUDY C. CORDER-BOLZ, BEVERLY DUNCAN,
JOHN FREEMAN, ERVING GOFFMAN, MICHAEL HINDELANG, ROSABETH
KANTER, ALAN C. KERCKHOFF, BARBARA LASLETT, PETER MANNING,
WILLIAM M. MASON, MURRAY MELBIN, ANTHONY OBERSCHALL, WHITNEY
POPE, SAMUEL PRESTON, DAVID SNYDER, BARBARA SOBIESZEK, AAGE
SØRENSEN, GERALD SUTTLES, DON H. ZIMMERMAN

Copy Editor: MARY S. MANDER

Managing Editor: LINNA McDADE

Editorial Assistant: CLYDENE MORGAN

Executive Officer: RUSSELL R. DYNES

SCOPE AND MISSION OF AMERICAN SOCIOLOGICAL REVIEW: The *American Sociological Review* publishes work of interest to the discipline in general, new theoretical developments, results of research that advances our understanding of the most fundamental social processes, and important methodological innovations. Like other publications of the Association, emphasis is given to exceptional quality. Unlike the more specialized journals of the Association, the *American Sociological Review's* primary focus is on whatever has the most general bearing on the knowledge of society.


Concerning manuscripts, address: Editor, *American Sociological Review*, Department of Sociology, University of Illinois, Urbana, Illinois 61801. (217) 333-6515.

The *American Sociological Review* is published at 49 Sheridan Avenue, Albany, New York, bimonthly in February, April, June, August, October, and December.

Applications for permission to quote from this journal should be addressed to the Executive Office.

Copyright © 1978 American Sociological Association.

ISSN 0003-1224



Annual membership dues of the Association: Member and International Member, \$30-50; Student Member, \$15; Associate, \$20; Student Associate, \$10; International Associate, \$12.

Application for membership and payment of dues should be made to the Executive Office. Subscription rate for members, \$10; nonmembers, \$15; institutions and libraries, \$30; nonmember students, \$10. Single issue, \$4.

New subscriptions and renewals will be entered on a calendar year basis only.

Change of address: Six weeks advance notice to the Executive Office and old address as well as new are necessary for change of subscriber's address.

Concerning advertising, changes of address, and subscriptions, address: Executive Office, American Sociological Association, 1722 N Street, N.W., Washington, D.C. 20036.

Claims for undelivered copies must be made within the month following the regular month of publication. The publishers will supply the missing copies when losses have been sustained in transit and when the reserve stock will permit.

Second class postage paid at Washington, D.C. and additional mailing offices.

Members of the Council

Officers of the Association for the year 1978

AMOS H. HAWLEY, *President*

North Carolina, Chapel Hill

ALICE S. ROSSI, *Vice President*

University of Massachusetts, Amherst

HUBERT BLALOCK, *President-Elect*

University of Washington

CHARLES Y. GLOCK, *Vice President-Elect*

University of California, Berkeley

JAMES F. SHORT, JR., *Secretary*

Washington State University

RUSSELL R. DYNES, *Executive Officer*

Ohio State University

J. MILTON YINGER, *Past President*

Oberlin College

Elected-at-Large

ELISE BOULDING

University of Colorado

ERNEST Q. CAMPBELL

Vanderbilt University

HERBERT L. COSTNER

National Science Foundation

RICHARD J. HILL

University of Oregon

JOAN HUBER

University of Illinois

LEWIS M. KILLIAN

University of Massachusetts

DAVID MECHANIC

University of Wisconsin

JOAN MOORE

University of Wisconsin

PAMELA A. ROBY

University of Cal., Santa Cruz

IMMANUEL WALLERSTEIN

SUNY, Binghamton

WILLIAM F. WHYTE

Cornell University

MAURICE ZEITLIN

University of Cal., Los Angeles

University of Wisconsin

(Articles in the REVIEW are indexed in the Social Sciences Index, Public Affairs Information Service, Psychological Abstracts, Sociological Abstracts, Ayer's Guide, University Microfilms, Abstracts for Social Workers, International Political Science Abstracts, and United States Political Science Documents.)

PROCESSING FEE

A processing fee of \$10 is required for each paper submitted; such fees will be waived for student members of ASA. This reflects a policy of the ASA Council and Committee on Publications affecting all ASA journals. It is a reluctant response to the rapidly accelerating costs of manuscript reviewing. A check or money order, made payable to the American Sociological Association should accompany each submission. The fee must be paid in order for the review process to begin.

ITEMS

■ We are sorry to lose the advice of Fred Davis who has resigned his position as Associate Editor. We are pleased that Michael Hindelang has agreed to come on board as Associate Editor for our 1978-1980 term.

■ JAMES R. KLUEGEL (*Causes and Cost of Racial Exclusion*) is Assistant Professor of Sociology at the University of California, Riverside. He is studying the relationship between organizational authority position and social and political beliefs. He also is pursuing research on the structure of American stratification belief systems.

■ ALEXANDER M. HICKS (*Class Power and State Policy*) is a Ph.D. Candidate in the Department of Sociology at the University of Wisconsin, Madison. His research interests include the political economy of governmental policy and legislative behavior at urban, state, and federal levels of government in the U.S. ROGER FRIEDLAND is Assistant Professor at the University of California, Santa Barbara. He does cross-national research on the political determinants of urban fiscal strains in the U.S., Germany, France, England, and Italy. He is author (with Robert Alford) of *Political Sociology* (Prentice-Hall, forthcoming) and editor of a special

(Continued on page 443)

AMERICAN SOCIOLOGICAL REVIEW

Official Journal of the American Sociological Association

AUGUST, 1978

VOLUME 43, NUMBER 4

ARTICLES

- Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature Thomas A. Heberlein and Robert Baumgartner 447
- A Cross-Cultural Study of Expressive and Instrumental Role Complementarity in the Family William D. Crano and Joel Aronoff 463
- Sex Differences in the Complexity of Children's Play and Games Janet Lever 471
- The Transition to Adulthood: Sex Differences in Educational Attainment and Age at Marriage Margaret Mooney Marini 483
- Institutionalization and Sick Role Identification among the Elderly John F. Myles 508
- Status Resources, Societal Reactions, and Type of Mental Hospital Admission William A. Rushing 521
- Is Status Attainment Research Atheoretical? Patrick M. Horan 534
- Nonrecursive Models of Labor Force Participation, Fertility Behavior, and Sex Role Attitudes Lynn Smith-Lovin and Ann R. Tickamyer 541
- A Rule for Inferring Individual-Level Relationships from Aggregate Data Glenn Firebaugh 557
- The Variable Order of Events in the Life Course Dennis P. Hogan 573

COMMENTS AND REPLIES

- Comment on Cortese, Falk and Cohen, ASR August, 1976
- On the Measurement of Segregation as a Random Variable Douglas S. Massey 587
- Understanding the Standardized Index of Dissimilarity: Reply to Massey Charles F. Cortese, R. Frank Falk and Jack Cohen 590

Editor: RITA J. SIMON

Deputy Editors: CLARK MCPHAIL AND ROSS M. STOLZENBERG

Associate Editors: MICHAEL T. AIKEN, ROBERT P. ALTHAUSER, PHILLIP F. BONACICH, ROBERT COLE, JUDY C. CORDER-BOLZ, BEVERLY DUNCAN, JOHN FREEMAN, ERVING GOFFMAN, MICHAEL HINDELANG, ROSABETH KANTER, ALAN C. KERCKHOFF, BARBARA LASLETT, PETER MANNING, WILLIAM M. MASON, MURRAY MELBIN, ANTHONY OBERSCHALL, WHITNEY POPE, SAMUEL PRESTON, DAVID SNYDER, BARBARA SOBIESZEK, AAGE SØRENSEN, GERALD SUTTLES, DON H. ZIMMERMAN

Copy Editor: MARY S. MANDER

Managing Editor: LINNA McDADE

Editorial Assistant: CLYDENE MORGAN

Executive Officer: RUSSELL R. DYNES

SCOPE AND MISSION OF AMERICAN SOCIOLOGICAL REVIEW: The *American Sociological Review* publishes work of interest to the discipline in general, new theoretical developments, results of research that advances our understanding of the most fundamental social processes, and important methodological innovations. Like other publications of the Association, emphasis is given to exceptional quality. Unlike the more specialized journals of the Association, the *American Sociological Review's* primary focus is on what ever has the most general bearing on the knowledge of society.

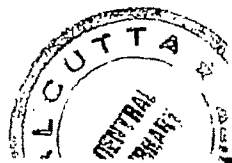
Concerning manuscripts, address: Editor, *American Sociological Review*, Department of Sociology, University of Illinois, Urbana, Illinois 61801. (217) 333-6515.

The *American Sociological Review* is published at 49 Sheridan Avenue, Albany, New York, bimonthly in February, April, June, August, October, and December.

Applications for permission to quote from this journal should be addressed to the Executive Office.

Copyright © 1978 American Sociological Association.

ISSN 0003-1224



Comment on Burgess and Nielsen, ASR June, 1974	
Threats to Internal Validity in Burgess and Nielsen's Social Exchange Experiment	592
..... W. David Pierce and W. Andrew Harrell	
Reply to Pierce and Harrell	597
..... Robert L. Burgess	
Comment on Chase-Dunn, ASR December, 1975	
The Effects of International Economic Dependence on Development: A Critique	600
..... Mark Stumpp, Robert M. Marsh, and Deborah A. Lake	
Reply to Stumpp, Marsh, and Lake	604
..... Christopher Chase-Dunn	
Reply to Walsh	605
..... J. Craig Jenkins	
Military Inequality: Reply to Hauser	607
..... John Sibley Butler	
Reply to Ronald L. Simons	610
..... Travis Hirschi and Michael J. Hindelang	
Reply to Stack	614
..... Richard Robinson	
Diamonds are Trumps: A Reply to House	615
..... Keith Hope	

Annual membership dues of the Association: Member and International Member, \$30-50; Student Member, \$15; Associate, \$20; Student Associate, \$10; International Associate, \$12.

Application for membership and payment of dues should be made to the Executive Office. Subscription rate for members, \$10; nonmembers, \$15; institutions and libraries, \$30; nonmember students, \$10. Single issue, \$4.

New subscriptions and renewals will be entered on a calendar year basis only.

Change of address: Six weeks advance notice to the Executive Office and old address as well as new are necessary for change of subscriber's address.

Concerning advertising, changes of address, and subscriptions, address: Executive Office, American Sociological Association, 1722 N Street, N.W., Washington, D.C. 20036.

Claims for undelivered copies must be made within the month following the regular month of publication. The publishers will supply the missing copies when losses have been sustained in transit and when the reserve stock will permit.

Second class postage paid at Washington, D.C. and additional mailing offices.

Members of the Council

Officers of the Association for the year 1978

AMOS H. HAWLEY, *President*

North Carolina, Chapel Hill

ALICE S. ROSSI, *Vice President*

University of Massachusetts, Amherst

HUBERT BLALOCK, *President-Elect*

University of Washington

CHARLES Y. GLOCK, *Vice President-Elect*

University of California, Berkeley

JAMES F. SHORT, JR., *Secretary*

Washington State University

RUSSELL R. DYNES, *Executive Officer*

Ohio State University

J. MILTON YINGER, *Past President*

Oberlin College

Elected-at-Large

ELISE BOULDING

University of Colorado

ERNEST Q. CAMPBELL

Vanderbilt University

HERBERT L. COSTNER

National Science Foundation

RICHARD J. HILL

University of Oregon

JOAN HUBER

University of Illinois

LEWIS M. KILLIAN

University of Massachusetts

DAVID MECHANIC

University of Wisconsin

JOAN MOORE

University of Wisconsin

PAMELA A. ROBY

University of Cal., Santa Cruz

IMMANUEL WALLERSTEIN

SUNY, Binghamton

WILLIAM F. WHYTE

Cornell University

MAURICE ZEITLIN

University of Cal., Los Angeles

University of Wisconsin

(Articles in the REVIEW are indexed in the Social Sciences Index, Public Affairs Information Service, Psychological Abstracts, Sociological Abstracts, Ayer's Guide, University Microfilms, Abstracts for Social Workers, International Political Science Abstracts, and United States Political Science Documents.)

ITEMS

Policy on Comments and Rejoinders

■ We shall endeavor to publish both a comment on an article and a rejoinder to a comment in the same issue. Comments should not exceed 2,500 words and rejoinders to comments are subject to the same length limits. All comments are sent out for review. Except under extraordinary circumstances, we will

not publish commentators' rebuttals to authors' replies. Comments are subject to the \$10.00 processing fee.

■ THOMAS A. HEBERLEIN (Factors Affecting Response Rates to Mailed Questionnaires) is Associate Professor in the Department of Rural Sociology at the University of Wisconsin, Madison. The major focus of his research involves social psychological

(Continued on page 620)

AMERICAN SOCIOLOGICAL REVIEW

Official Journal of the American Sociological Association

OCTOBER, 1978

VOLUME 43, NUMBER 5

ARTICLES

The Hawthorn Experiments: First Statistical Interpretation	Richard Herbert Franke and James D. Kaul	623
The Myth of Social Class and Criminality: An Empirical Assessment of the Empirical Evidence	Charles R. Tittle, Wayne J. Vilemez, and Douglas A. Smith	643
Sex Differences in Soviet Earnings	Michael Swafford	657
The Influence of "Family Background" on Intellectual Attainment	Sandra Scarr and Richard A. Weinberg	674
Cross Pressures during Socialization for Medicine	Israel Adler and Judith T. Shuval	693
Stratification in a Dual Economy: A Sectoral Model of Earnings Determination	E. M. Beck, Patrick M. Horan, and Charles M. Tolbert II	704
Power, Equity and Commitment in Exchange Networks	Karen S. Cook and Richard M. Emerson	721
Cross-National Diffusion: The Effects of Canadian TV on Rural Minnesota Viewers	David E. Payne	740
Black Political Participation in the United States: Some Recent Evidence	Nicholas L. Danigelis	756

COMMENT AND REPLY

Comment on Szymanski, ASR June, 1976		
Black Subordination and White Economic Well-Being	Wayne J. Vilemez	772
White Workers' Loss from Racial Discrimination: A Reply to Vilemez	Albert Szymanski	776

Editor: RITA J. SIMON

Deputy Editors: CLARK McPHAIL AND ROSS M. STOLZENBERG

Associate Editors: MICHAEL T. AIKEN, ROBERT P. ALTHAUSER, PHILLIP F. BONACICH, ROBERT COLE, JUDY C. CORDER-BOLZ, BEVERLY DUNCAN, JOHN FREEMAN, ERVING GOFFMAN, MICHAEL HINDELANG, ROSABETH KANTER, ALAN C. KERCKHOFF, BARBARA LASLETT, PETER MANNING, WILLIAM M. MASON, MURRAY MELBIN, ANTHONY OBERSCHALL, WHITNEY POPE, SAMUEL PRESTON, DAVID SNYDER, BARBARA SOBIESZEK, AAGE SØRENSEN, GERALD SUTTLES, DON H. ZIMMERMAN

Copy Editor: MARY S. MANDER

Managing Editor: LINNA McDADE

Editorial Assistant: CLYDENE MORGAN

Executive Officer: RUSSELL R. DYNES

SCOPE AND MISSION OF AMERICAN SOCIOLOGICAL REVIEW: The *American Sociological Review* publishes work of interest to the discipline in general, new theoretical developments, results of research that advances our understanding of the most fundamental social processes, and important methodological innovations. Like other publications of the Association, emphasis is given to exceptional quality. Unlike the more specialized journals of the Association, the *American Sociological Review's* primary focus is on whatever has the most general bearing on the knowledge of society.


Concerning manuscripts, address: Editor, *American Sociological Review*, Department of Sociology, University of Illinois, Urbana, Illinois 61801. (217) 333-6515.

The *American Sociological Review* is published at 49 Sheridan Avenue, Albany, New York, bimonthly in February, April, June, August, October, and December.

Applications for permission to quote from this journal should be addressed to the Executive Office.

Copyright © 1978 American Sociological Association.

ISSN 0003-1224



Annual membership dues of the Association: Member and International Member, \$30-50; Student Member, \$15; Associate, \$20; Student Associate, \$10; International Associate, \$12.

Application for membership and payment of dues should be made to the Executive Office. Subscription rate for members, \$10; nonmembers, \$15; institutions and libraries, \$30; nonmember students, \$10. Single issue, \$4.

New subscriptions and renewals will be entered on a calendar year basis only.

Change of address: Six weeks advance notice to the Executive Office and old address as well as new are necessary for change of subscriber's address.

Concerning advertising, changes of address, and subscriptions, address: Executive Office, American Sociological Association, 1722 N Street, N.W., Washington, D.C. 20036.

Claims for undelivered copies must be made within the month following the regular month of publication. The publishers will supply the missing copies when losses have been sustained in transit and when the reserve stock will permit.

Second class postage paid at Washington, D.C. and additional mailing offices.

Members of the Council

Officers of the Association for the year 1978

AMOS H. HAWLEY, *President*

North Carolina, Chapel Hill

ALICE S. ROSSI, *Vice President*

University of Massachusetts, Amherst

HUBERT BLALOCK, *President-Elect*

University of Washington

CHARLES Y. GLOCK, *Vice President-Elect*

University of California, Berkeley

JAMES F. SHORT, JR., *Secretary*

Washington State University

RUSSELL R. DYNES, *Executive Officer*

Ohio State University

J. MILTON YINGER, *Past President*

Oberlin College

Elected-at-Large

ELISE BOULDING

University of Colorado

ERNEST Q. CAMPBELL

Vanderbilt University

HERBERT L. COSTNER

National Science Foundation

RICHARD J. HILL

University of Oregon

JOAN HUBER

University of Illinois

LEWIS M. KILLIAN

University of Massachusetts

DAVID MECHANIC

University of Wisconsin

JOAN MOORE

University of Wisconsin

PAMELA A. ROBY

University of Cal., Santa Cruz

IMMANUEL WALLERSTEIN

SUNY, Binghamton

WILLIAM F. WHYTE

Cornell University

MAURICE ZEITLIN

University of Cal., Los Angeles

University of Wisconsin

(Articles in the REVIEW are indexed in the Social Sciences Index, Public Affairs Information Service, Psychological Abstracts, Sociological Abstracts, Ayer's Guide, University Microfilms, Abstracts for Social Workers, International Political Science Abstracts, and United States Political Science Documents.)

New Policy

- Starting as soon as possible, the ASR will introduce a Research Notes and Reports Section. We will be interested in receiving manuscripts that fit those qualifications.
- On the regular articles, we are loosening our page limitations somewhat, and would be happy to review manuscripts that are up to fifty manuscript pages.

ITEMS

■ RICHARD HERBERT FRANKE (The Hawthorne Experiments) is Associate Professor in the Department of Management, Worcester Polytechnic Institute. His studies include the social factors in industrial performance, the effects of industrial pollutants on illness, and the determinants of crime.

JAMES D. KAUL earned his M.S. degree in Urban Planning at the University of Wisconsin, Milwaukee, and is working on the social, economic, and health impacts of air pollution for the U.S. Environmental Protection Agency at the Indiana Board of Health in Indianapolis. He is completing (with Richard Franke) a historical review of the human relations movement.

(Continued on page 783)

AMERICAN SOCIOLOGICAL REVIEW

Official Journal of the American Sociological Association

DECEMBER, 1978

VOLUME 43, NUMBER 6

ARTICLES

- Presidential Address: Cumulative Change in Theory and in History Amos H. Hawley 787
The Ideology of Childhood and the State: Rules Distinguishing Children in National Constitu-
tions, 1870-1970 John Boli-Bennett and John W. Meyer 797
Bringing the Boss Back In: Employer Size, Employee Schooling, and Socioeconomic
Achievement Ross M. Stolzenberg 813
Cultural Effects on Organizational Structure: The Case of Japanese Firms in the United
States James R. Lincoln, Jon Olson and Mitsuyo Hanada 829
White Enrollment Stability and School Desegregation: A Two-Level Analysis
..... Micheal W. Giles 848
Measures of Inequality Paul D. Allison 865
The Effect of Direct Government Involvement in the Economy on the Degree of Income
Inequality: A Cross-National Study Steven Stack 880
Productivity and Academic Position in the Scientific Career J. Scott Long 889
Occupational Sex Identification and the Assessment of Male and Female Earnings Inequality
..... Steven D. McLaughlin 909

COMMENTS AND REPLIES

- Comment on Hagan and Leon, ASR August, 1977
Marxist and "Organizational" Approaches to Delinquency and the Sociology of Law: Crucial
Problems in "Testing" the Perspectives Pat Lauderdale and Gerald Larson 922
Crass Conflict: A Reply to Lauderdale and Larson John Hagan and Jeffrey Léon 925
Comment on Melbin, ASR February, 1978
On the Use of Historical Documents and Evidence Burton B. Silver 929
Reply to Silver: On the Structure of Evidence Murray Melbin 931

Editor: RITA J. SIMON

Deputy Editors: CLARK MCPHAIL AND ROSS M. STOLZENBERG

Associate Editors: MICHAEL T. AIKEN, ROBERT P. ALTHAUSER, PHILLIP F.
BONACICH, ROBERT COLE, JUDY C. CORDER-BOLZ, BEVERLY DUNCAN,
JOHN FREEMAN, ERVING GOFFMAN, MICHAEL HINDELANG, ROSABETH
KANTER, ALAN C. KERCKHOFF, BARBARA LASLETT, PETER MANNING,
WILLIAM M. MASON, MURRAY MELBIN, ANTHONY OBERSCHALL, WHITNEY
POPE, SAMUEL PRESTON, DAVID SNYDER, BARBARA SOBIESZEK, AAGE
SØRENSEN, GERALD SUTTLES, DON H. ZIMMERMAN

Copy Editor: MARY S. MANDER

Managing Editor: LINNA MCDADE

Editorial Assistant: CLYDENE MORGAN

Executive Officer: RUSSELL R. DYNES

SCOPE AND MISSION OF AMERICAN SOCIOLOGICAL REVIEW: The *American Sociological Review* publishes work of interest to the discipline in general, new theoretical developments, results of research that advances our understanding of the most fundamental social processes, and important methodological innovations. Like other publications of the Association, emphasis is given to exceptional quality. Unlike the more specialized journals of the Association, the *American Sociological Review's* primary focus is on what-
ever has the most general bearing on the knowledge of society.

Concerning manuscripts, address: Editor, *American Sociological Review*, Department of
Sociology, University of Illinois, Urbana, Illinois 61801. (217) 333-6515.

The *American Sociological Review* is published at 49 Sheridan Avenue, Albany, New York,
bimonthly in February, April, June, August, October, and December.

Applications for permission to quote from this journal should be addressed to the Executive
Office.

Copyright © 1978 American Sociological Association.

ISSN 0003-1224



Annual membership dues of the Association: Member and International Member, \$30–50; Student Member, \$15; Associate, \$20; Student Associate, \$10; International Associate, \$12.

Application for membership and payment of dues should be made to the Executive Office. Subscription rate for members, \$10; nonmembers, \$15; institutions and libraries, \$30; nonmember students, \$10. Single issue, \$4.

New subscriptions and renewals will be entered on a calendar year basis only.

Change of address: Six weeks advance notice to the Executive Office and old address as well as new are necessary for change of subscriber's address.

Concerning advertising, changes of address, and subscriptions, address: Executive Office, American Sociological Association, 1722 N Street, N.W., Washington, D.C. 20036.

Claims for undelivered copies must be made within the month following the regular month of publication. The publishers will supply the missing copies when losses have been sustained in transit and when the reserve stock will permit.

Second class postage paid at Washington, D.C. and additional mailing offices.

Members of the Council

Officers of the Association for the year 1978

AMOS H. HAWLEY, *President*

North Carolina, Chapel Hill

ALICE S. ROSSI, *Vice President*

University of Massachusetts, Amherst

HUBERT BLALOCK, *President-Elect*

University of Washington

CHARLES Y. GLOCK, *Vice President-Elect*

University of California, Berkeley

JAMES F. SHORT, JR., *Secretary*

Washington State University

RUSSELL R. DYNES, *Executive Officer*

Ohio State University

J. MILTON YINGER, *Past President*

Oberlin College

Elected-at-Large

ELISE BOULDING

University of Colorado

ERNEST Q. CAMPBELL

Vanderbilt University

HERBERT L. COSTNER

National Science Foundation

RICHARD J. HILL

University of Oregon

JOAN HUBER

University of Illinois

LEWIS M. KILLIAN

University of Massachusetts

DAVID MECHANIC

University of Wisconsin

JOAN MOORE

University of Wisconsin

PAMELA A. ROBY

University of Cal., Santa Cruz

IMMANUEL WALLERSTEIN

SUNY, Binghamton

WILLIAM F. WHYTE

Cornell University

MAURICE ZEITLIN

University of Cal., Los Angeles

University of Wisconsin

(Articles in the REVIEW are indexed in the Social Sciences Index, Public Affairs Information Service, Psychological Abstracts, Sociological Abstracts, Ayer's Guide, University Microfilms, Abstracts for Social Workers, International Political Science Abstracts, and United States Political Science Documents.)

New Policy

- Starting as soon as possible, the ASR will introduce a Research Notes and Reports Section. We will be interested in receiving manuscripts that fit those qualifications.
- On the regular articles, we are loosening our page limitations somewhat, and would be happy to review manuscripts that are up to fifty manuscript pages.

ITEMS

■ Because this issue marks the end of my first year as editor, it is a good time to express my appreciation to Clark McPhail and Ross Stolzenberg for all of the support and good advice they have given me. Both of them will continue on as Deputy Editors. I want, also, to thank the Associate Editors who have com-

pleted their three-year terms and are stepping down: Robert P. Althaus, Phillip Bonacich, John Freeman, Erving Goffman, Alan Kerckhoff, Peter Manning, Whitney Pope, Aage Sørensen, Gerald Suttles, Judy Corder-Bolz and Don H. Zimmerman. Beginning January 1, 1979 the following persons have been asked and have agreed to work with us as

(Continued on page 933)

AMERICAN SOCIOLOGICAL REVIEW

February 1978

Volume 43 Number 1

NIGHT AS FRONTIER*

MURRAY MELBIN

Boston University

American Sociological Review 1978, Vol. 43 (February):3-22

While the settlement of some of the world's land areas was coming to an end, there began an increase in wakeful activity over more of the 24-hour day. This trend of expansion in time is continuing, especially in urban areas. The hypothesis that night has become the new frontier is supported by the premise that time, like space, can be occupied and is treated so by humans. A set of evidence, including results of several field experiments, show that nighttime social life in urban areas resembles social life on former land frontiers. The research data refers mainly to contemporary Boston and to the U.S. West a century ago.

Humans are showing a trend toward more and more wakeful activity at all hours of day and night. The activities are extremely varied. Large numbers of people are involved. And the trend is worldwide. A unifying hypothesis to account for it is that night is a frontier, that expansion into the dark hours is a continuation of the geographic migration across the face of the earth. To support this view, I will document the trend and then offer a premise about the nature of time and its relation to space. Third, I will show that social life in the nighttime has many important characteristics that resemble social life on land frontiers.

THE COURSE OF EXPANSION

We were once a diurnal species bounded by dawn and dusk in our wakeful

activity. Upon mastering fire, early humans used it for cooking and also for sociable assemblies that lasted for a few hours after darkness fell. Some bustle throughout the 24-hour cycle occurred too. Over the centuries there have been fires tended in military encampments, prayer vigils in temples, midnight betrothal ceremonies, sentinels on guard duty at city gates, officer watches on ships, the curing ceremonies of Venezuelan Indians that begin at sundown and end at sunrise, innkeepers serving travelers at all hours. In the first century A.D., Rome was obliged to relieve its congestion by restricting chariot traffic to the night hours (Mumford, 1961:217).

Yet around-the-clock activity used to be a small part of the whole until the nineteenth century. Then the pace and scope of wakefulness at all hours increased smartly. William Murdock developed a feasible method of coal-gas illumination and, in 1803, arranged for the interior of the Soho works in Birmingham, England to be lighted that way. Other mills nearby began to use gas lighting. Methods of distributing coal-gas to all buildings and street lamps in a town were introduced soon after. In 1820 Pall Mall in London became the first street to be lit by coal-gas. Artificial lighting gave great stimulus to the nighttime entertainment industry (Schlesinger, 1933:105). It also permitted multiple-shift factory opera-

* I thank the Center for Studies of Metropolitan Problems, National Institute of Mental Health, for grant MH-22763 through which the research and the preparation of this essay was supported; and Earl Mellor of the Bureau of Labor Statistics for providing interpretive help and data tables from the 1976 Current Population Survey. I also thank my research assistants William O. Clarke, Ann Getman, Shelley Leavitt, Lee Parmenter, Alan Rubenstein, Melanie Wallace, and Marilyn Arsem for field observations at all hours in rain and bitter cold as well as mild weather; and my colleagues Paul Hollander and Anthony Harris of the University of Massachusetts at Amherst, for serving as recipients in the lost-key test.

tions on a broad scale. Indeed by 1867 Karl Marx (1867:chap. 10, sec. 4) was to declare that night work was a new mode of exploiting human labor.

In the closing decades of the nineteenth century two developments marked the changeover from space to time as the realm of human migration in the United States. In 1890 the Bureau of the Census announced that the land frontier in America had come to an end, for it was no longer possible to draw a continuous line across the map of the West to define the edge of farthest advance settlement. Meanwhile, the search for an optimum material for lantern lights, capable of being repeatedly brought to a white heat, culminated in 1885 in the invention of the Welsbach mantle—a chemically impregnated cotton mesh. The use of the dark hours increased thereafter, and grew further with the introduction of electric lighting.

Here and there one may find documentation of the trend. During the First World War there was selective concern, expressed by Brandeis and Goldmark (1918) in *The Case Against Night Work for Women*, about the impact of off-hours work. A decade later the National Indus-

trial Conference Board (1927) published a comprehensive survey with an account of the characteristics of the off-hours workers.

The most systematic evidence of steadily increasing 24-hour activity in the U.S. is the growth of radio and television broadcasting. Broadcasters authorize surveys to learn about the market that can be reached in order to plan programs and to set advertising rates. The number of stations active at given hours and the spread of those hours around the clock reflects these research estimates of the size of the wakeful population—the potential listeners. Table 1 shows trends in the daily schedule spanning the entire periods of commercial broadcasting for both radio and television. Although not shown in the table, television hours in Boston ended at 11:30 p.m. in 1949, and then widened to include the Late Show and then the Late Late Show in the intervening years until 1974. Each medium has moved increasingly to 24-hour programming and mirrors the growth in nighttime activity.

In the present decade, for the first time, the U.S. Bureau of Labor Statistics (1976: Table 1) asked about the times of day that people worked. In 1976, of 75 million in

Table 1. Numbers of Radio and Television Stations and Their Hours of Broadcasting in Boston^a

	The Span of Commercial Broadcasting									
	April 1929	April 1934	April 1939	April 1944	April 1949	April 1954	April 1959	April 1964	April 1969	April 1974
Radio										
Number of stations	7	7	8	7	8	14	15	20	26	27
Number of 24-hour stations	0	0	0	0	0	1	3	8	12	15
Percent of 24-hour stations						7%	20%	40%	46%	57%
Television										
Number of stations					2	4	4	4	5	7
Number of 24-hour stations					0	0	0	0	0	1
Percent of 24-hour stations										14%

^a Sources: listings in Boston newspapers—*Globe*, *Herald*, *Record*, and *Traveler*—and the broadcasters themselves. If the content of a broadcaster's AM and FM radio programming or VHF and UHF television programming differs, that broadcast is counted as two stations.

the work force, 12 million reported they were on the job mainly after dark and 2.5 million of those persons worked a full shift beginning about midnight. Since these figures do not include *the clientele* that used such establishments as restaurants, hospital emergency wards, gambling rooms, and public transportation, these numbers are conservative estimates of how many people are up and about at night.

Today more people than ever are active outside their homes at all hours engaged in all sorts of activities. There are all-night supermarkets, bowling alleys, department stores, restaurants, cinemas, auto repair shops, taxi services, bus and airline terminals, radio and television broadcasting, rent-a-car agencies, gasoline stations. There are continuous-process refining plants, and three-shift factories, post offices, newspaper offices, hotels, and hospitals. There is unrelenting provision of some utilities—electric supply, staffed turnpike toll booths, police patrolling, and telephone service. There are many emergency and repair services on-call: fire fighters, auto towing, locksmiths, suppliers of clean diapers, ambulances, bail bondsmen, insect exterminators, television repairers, plate glass installers, and funeral homes.

The trend of nighttime expansion is under way outside the United States as well. In Great Britain since the Second World War, the yearly increase in the percentage of the manual labor force on shifts in manufacturing has been about 1% a year, and greater increases have been noted in vehicle manufacture and in the chemical industry (Young and Willmott, 1973:175). Meier (1976:965) observes that Singapore is becoming one of the most intensive 24-hour cities. Data on around-the-clock activity in Peru, France, U.S.S.R. and eight other nations is provided in a volume on *The Use of Time* (Szalai, 1972:appendices).

SPACE AND TIME FRONTIERS AND SETTLEMENTS

Time, like space, is part of the ecological niche occupied by a species. Although every type exists throughout the 24-hour cycle, to reflect the way a species uses its

niche we label it by *the timing of its wakeful life*. The terms diurnal and nocturnal refer to the periods the creatures are active. We improve our grasp of the ecology of a region by recognizing the nighttime activity of raccoons, owls and rats, as well as by knowing the spatial dispersion of these and other animals. The same area of a forest or meadow or coral reef is used incessantly, with diurnal and nocturnal creatures taking their active turns. We make geographic references to humans in a similar way. We refer to an island people or a desert people, or the people of arctic lands as a means of pointing out salient features of their habitats.

This similar treatment of time and space rests on the assumption that both of them are containers for living. Consider the dictionary definition of the word *occupy*: "2. To fill up (take time or space): a lecture that occupied three hours" (*American Heritage Dictionary*, 1970:908). Geographers study activities rather than physical structures to decide whether and how people occupy space (Buttimer, 1976:286). The mere presence of buildings and related physical structures in places like Machu-Pichu, Petra, and Zimbabwe do not make us believe they are habitations now. The once-boisterous mining centers in the American West that have become ghost towns are settlements no longer. Conversely, we say a farming region in which people are active is inhabited even though buildings are few. The presence of human-built structures is not the criterion for occupying a region, it is people and their activities.

Like rural settlements, the occupation of time need not be dense. For example, London Transport lists 21 all-night bus routes. On many of these routes "all-night" service means no more than once an hour. Yet, even though the bus does not pass during the intervening 59 minutes, the schedule is said to be continuous. If an active moment interacts with quiet moments around it, the entire period is taken as occupied.

Of course, no time has ever been used without also using it in some place. No space has ever been used without also using it some hours of the day. Space and time together form the container of life

activity. We forget this in the case of former frontiers because expansion then occurred so dramatically across the land. Less notice was paid to the 16 hours of wakefulness because the daily use of time was rather constant as the surge of geographic expansion kept on over the face of the earth. As time use remained unchanged, it was disregarded in human ecological theory. In different eras, however, expansion may proceed more rapidly in either space or time. Recently expansion is taking place in time. Since people may exploit a niche by distributing themselves and their activities over more hours of the day just as they do by dispersing in space, a frontier could occur in the time dimension too.

A *settlement* is a stable occupation of space and time by people and their activities. A *frontier* is a pattern of sparse settlement in space or time, located between a more densely settled and a practically empty region. Below a certain density of active people, a given space-time region is a wilderness. Above that point and continuing to a higher level of density, the presence of people in activities will make that area a frontier. Above that second cutoff point the further denseness of active people turns the area into a fully inhabited region. In a given historical period the frontier's boundaries may be stable or expanding. When expanding the frontier takes on the aspect of venturing into the unknown and is often accompanied by novelty and change.

SIMILARITIES BETWEEN LAND FRONTIERS AND TIME FRONTIERS

Two kinds of evidence would support the hypothesis of night as frontier. One is that the forces for expansion into the dark hours are the same as those resulting in expansion across the land. That is, a single causal explanation should account for the spread of people and their activities, whether in space or in time. I offered such an outline in another essay; it includes enabling factors, demand push, supply pull, and stabilizing feedback (Melbin, 1977). The other line of evidence is that the same important features of social life should be found both in time and

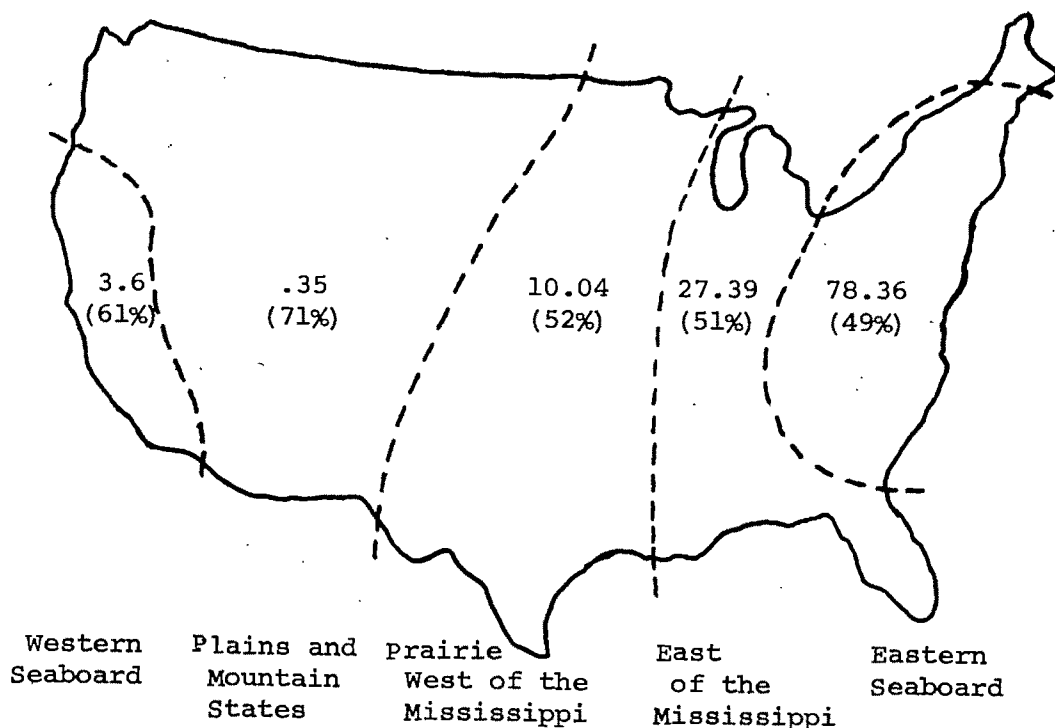
in space frontiers. The rapid expansion in after-dark activity has been taking place mostly in urban areas. Therefore the culture of the contemporary urban nighttime should reveal the same patterns and moods found in former land frontiers.

I have chosen to review life in the U.S. West in the middle of the nineteenth century along with the present-day nighttime. Of course there were other land frontiers and the hypothesis should apply to all of them. However there are good reasons to begin by demonstrating it for the U.S. West. One is that the archives holding information about this westward flow are thorough, well organized, and readily available. Another reason is that the U.S. West has continuity with expansion into the night. The movement westward reached the California coast. California's main cities have since become areas of great activity in the dark hours, as if the flow across the continent swerved into the nighttime rather than spilling into the sea.

Specifically, the land frontier to be discussed is the area west of the Mississippi River during the middle decades of the nineteenth century, about 1830–1880. The urban nighttime will be any major urban area during the stretch from about midnight to 7:30 a.m. during the decades of the 1960s and 1970s. Most of my examples will be findings from a recent study of Boston. There are many aspects in which social life at night is like the social life of other frontiers.

1. *Advance Is in Stages*

There is a succession of steps in colonizing any new region. People ventured into the western outskirts "in a series of waves . . . the hunter and the fur trader who pushed into the Indian country were followed by the cattle raiser and he by the pioneer farmer" (Turner, 1965:59; 1893:12, 19–20). Life styles were distinctive in each stage as well. The hunters and trappers did not dwell like the miners who followed, and they in turn lived differently from the pioneer farmers who came later (Billington, 1949:4–5). Although living conditions were generally crude then, there was a decided increase in comfort



• Source: U.S. Bureau of the Census, 1975: Vol. 1, ser. A195-209.

Figure 1. Persons per Square Mile (and Percent Males per Square Mile) during the 1870-1880 Decade*

for the farmers settled in one place compared with the earlier-day trappers who were usually on the move.

There is also a succession of phases in settling the nighttime. Each stage fills the night more densely than before and uses those hours in a different way. First came isolated wanderers on the streets; then groups involved in production activities, the graveyard-shift workers. Still later those involved in consumption activities arrived, the patrons of all-night restaurants and bars, and the gamblers who now cluster regularly by midnight at the gaming table in resorts.

The rates of advance are unequal in both cases. Population gains and development are not unbroken. In the West economic growth was erratic. Periods of depression, dry seasons and other hardships drove many people to abandon their homesteads and move back east. Similarly, during the oil embargo of 1973-1974 there was some retreat from nighttime activity, as restaurants and auto service stations and other businesses cut back hours of serving the public.

2. Population Is Sparse and Also More Homogenous

At first only a few people venture into the new region. The frontier line in the U.S. West was drawn by the Census Bureau through an area of density of two to six inhabitants per square mile. The other side of the line was tabbed the "wilderness." The demographic composition of the western frontier was mostly vigorous young males with proportionately fewer females and aged persons than found in the populations of the eastern states (Riegel, 1947:624; Godkin, 1896:13; Dick, 1937:7, 232). This demographic picture fits the night as well. There are fewer people up and about and most of them are young males.

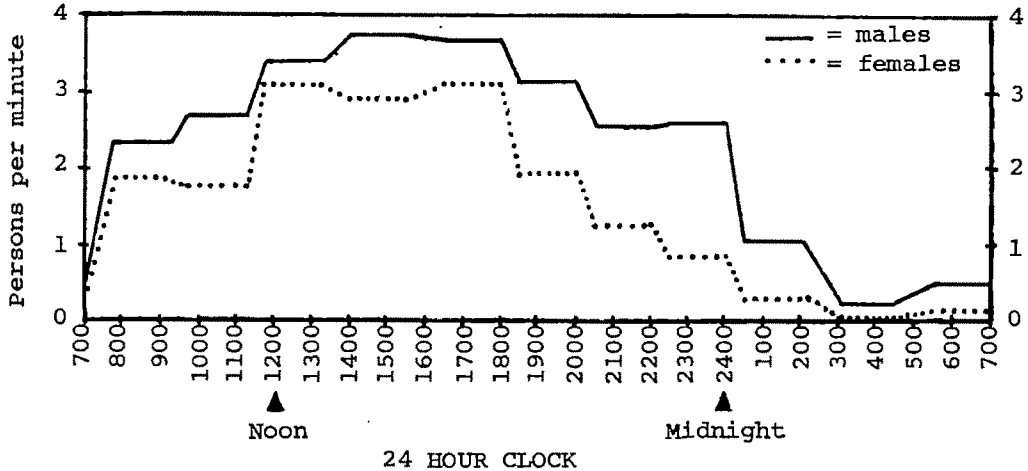
A crude comparison between a frontier line in the nineteenth-century West and a time interval in the twentieth-century nighttime is possible if the map of Figure 1 is scanned from right to left and the graph of Figure 2 is scanned from left to right. In this view both figures show similar graded densities. In Figure 2, the period after

Total
per
minute

4.12	4.52	6.50	6.71	6.85	5.11	3.85	3.46	1.32	.28	.59
------	------	------	------	------	------	------	------	------	-----	-----

% Male

56%	61%	52%	56%	54%	61%	67%	75%	78%	89%	75%
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----



^a The tallies were accomplished during field visits based on a stratified random sample in the city of Boston in 1974. Along with this person-tally, other measurements including the experiments described later in this section were carried out during these field visits. The 166 field visits (56 in early June, 56 in early September, and 54 in early December), each two hours long, were distributed among three sites: a shopping street, a residential street, and a transportation hub (62, 42, and 62 visits, respectively), each located one mile from the next along a line through the center city. In all, we made 456 five-minute tallies of people passing one-way at eight checkpoints.

The times of field visits were randomly selected from two-hour intervals around the 24-hour clock. I established 11 such intervals in this sampling frame for two reasons: certain times of day are clearly associated with certain kinds of activities, especially 0730–0929 as the morning rush hours and 1615–1814 as the evening rush hours. Given the two rush-hour spans placed as they are in the 24-hour period, and given the need for two-hour field visits in order to accomplish the experiments and other studies in the more comprehensive program of investigating activity around the clock, it was impossible to fit the remaining hours of the day neatly into two-hour intervals throughout the period. To insert 12 two-hour intervals into the period would force some the strata to be out of phase with familiar time intervals in Boston. So 11 strata were established and the remaining scattered minutes left unsampled, as follows: morning rush hour = 0730–0929; daytime = 0930–1129, 1130–1329 (1330–1344 omitted), 1345–1544 (1545–1614 omitted); evening rush hour = 1615–1814; evening = 1815–2014, 2015–2214, 2215–0014; and night = 0015–0214 (0215–0244 omitted), 0245–0444 (0445–0514 omitted), 0515–0714 (0715–0729 omitted). Thus the two longest phases (day and night) lose the unsampled minutes, 45 minutes from the daytime and 75 minutes at night. By removing these amounts from the sample, the day, evening and night phases are six hours each and the rush hours are two hours each. (It may be easier to design the sample for 12 two-hour periods and then align the boundaries of each time stratum to local patterns in the analysis stage of the research. If this is done, records of measurement should be made to the minute.) The number of visits for the five phases of the day were: morning rush hour, 13; day, 47; evening rush hour, 11; evening, 48; and night, 47. The number counts for each time stratum were divided by the number of tally minutes for that stratum to equalize the sampling fractions across strata, and then doubled to provide two-way estimates.

Three teams of male and female researchers, with pairings reshuffled systematically across times of day and across sites, carried out these tallies and the experiments reported later in this section.

Figure 2. Females and Males Passing per Minute on the Streets of Central Boston^a

midnight until 7 a.m. is sparsest and stands in the same relation to the rest of the day as the region west of the Mississippi stands in relation to the East in Figure 1. The figures also show that the proportion of males in the population is

higher on the frontiers. Just as this part of the total is largest in the Plains and Mountain States (71%), males comprise the largest part of the street population (89%) in the middle of the night.

Estimates of the ages of passersby were

also made during the field observations that yielded the data for Figure 2.¹ Whereas people of all ages were on the streets during the day, no one over 59 was seen between midnight and 5 a.m.; and from 2 to 5 a.m. no one over 41 was seen.

3. There Is Welcome Solitude, Fewer Social Constraints, and Less Persecution

The land frontier offered tranquillity, a place for relief from feelings of being hemmed in. "Fur traders . . . were psychological types who found forest solitudes more acceptable than the company of their fellow men" (Billington, 1949:4). It was appealing to escape into the wilderness, to leave deceit and disturbance, and vexing duties and impositions of the government behind (Robbins, 1960:148). "'Oh, how sweet,' wrote William Penn from his forest refuge, 'is the quiet of these parts, freed from the troubles and perplexities of woeful Europe'" (Turner, 1893:262). Even later the West was "a refuge . . . from the subordination of youth to age" (Turner, 1932:25). The outer fringes offered escape from persecution too. Mormons and Hutterites both made their ways westward to avoid harassment from others.

In a parallel way, many have enjoyed the experience of walking at night along a street that is ordinarily jammed during the day. Individuals who are up and about then report a feeling of relief from the crush and anonymity of daytime city life. The calm of those hours is especially appealing to young people, who come to feel that they possess the streets. (A test of this proposition must of course control for the fear of criminal assault in the dark; I will discuss this further in items 7 and 8 below.) Also, a portion of the people out at night are those avoiding social constraints and perhaps persecution. Street people and homosexuals, for example,

find more peace in the dark because surveillance declines. Some night owls are urban hermits. Some individuals who are troubled or stigmatized—such as the very ugly or obese—retreat from the daytime to avoid humiliation and challenge. They stay up later, come out when most others are gone, and are more secure as they hobnob with nighttime newsdealers and porters and elevator men. In this way the night affords an outlet. Like the West it serves an insulating function that averts possible tensions from unwanted encounters.

4. Settlements Are Isolated

Initially migration beyond the society's active perimeter is scattered. The land frontier settlements were small and apart from one another. There was little communication across districts and much went on in each in a self-sufficient way. People in the East did not think of the relevance of borderland activities for their own existence and the pioneers were indifferent to outside society (Billington, 1949:96, 746).

As the city moves through phases of the day it switches from coordinated actions to unconnected ones. Pockets of wakeful activity are separated from one another, are small scale compared to daytime events, and there is less communication between the pockets. The people of the daytime give little thought to those active in the dark and do not view them as part of the main community.

5. Government Is Initially Decentralized

Whatever high-level group may decide the laws and policies for a nation or a community, outside the purview of superiors there are subordinates who make decisions that would otherwise be the domain of the higher-ups or subject to their approval. As the land frontier moved farther from the national center of policy making, the interpretation of the law and judicial decisions were carried out by individuals who were rarely checked on and who rarely consulted with their superiors. Hollon (1973:96) notes that events took place "remote from the courts

¹ A comparison of the age estimate made by an observer and the answer to an age query made of 696 passersby yielded a correlation (within two years) of .96 for the six observers, with the lowest coefficient for an observer being .93. Populations at these sites are somewhat younger than the city's census average.

of authorities . . . [and] the frontiersmen not only enforced their own law, they chose which laws should be enforced and which should be ignored."

Today, although many organizations and cities are continually active, their primary administrators—directors, heads of departments, mayors—are generally on duty only during the daytime. At night they go to sleep and a similar decentralization of power follows. To some extent this is an explicit delegation of authority. But discretion is stretched for other reasons too. Night nurses decide not to wake up the doctor on duty because he gets annoyed at being disturbed for minor problems (Kozak, 1974:59). Shift supervisors choose not to bother the plant manager for similar reasons. Lesser officials make decisions that in the daytime are left for higher-ranking administrators. The style and content of the way the organization or the city is run at night changes accordingly. For example, for the same types of cases, decisions by police officers at night will be based less on professional role criteria and more on personal styles. This results in more extreme instances of being strict and lenient, arbitrary and humane.

6. *New Behavioral Styles Emerge*

Both land and time frontiers show more individualism because they are remote, the environment is unusual (compared with the centers of society), and others subjected to the same conditions are tolerant. Those who traveled to the western borders broke from ordinary society. The casual observance by others, the constituted authority, and the familiar settings and the norms they implied were gone. This left room for unconventional behavior. Easterners thought westerners were unsavory. The president of Yale College said, "The class of pioneers cannot live in regular society. They are too idle, too talkative, too passionate, too prodigal, and too shiftless to acquire either property or character" (cited in Turner, 1893:251). Another traveler in the same period wrote, "It is true there are worthless people here [in settlements hundreds of miles from any court of justice] and the most so, it must be confessed, are from

New England" (Flint, 1826:402). He did go on to say that there were also many who were worthy.

Deviance was also *created* out west. Many pioneer wives lived on the plains for extended periods without ordinary social contacts, especially when their husbands left on journeys for days or weeks. These women often became withdrawn and untalkative, so shy and uneasy with strangers that they would run away when one approached (Humphrey, 1931:128). From the evidence at hand, these were normal, happy women in the cities when they were growing up, but they were affected by the frontier environment. On the western boundary people were used to this behavior on the part of lonely, isolated women and accepted it. In the eastern cities the same conduct would have been taken as odd.

There is also a popular image of the night as the haunt of weirdos and strange characters, as revealed in comments like "I don't know where they hide during the day but they sure come out after dark." Moreover, at night one can find people who, having lived normal lives, are exposed to unusual circumstances that draw them into unconventional behavior. Becker (1963:79, 97, 98) gives such an account of jazz musicians. They work late in the evening and then associate with very few daytime types in their recreation after midnight. The milieu harbors a deviant subculture that is tolerated and even expected.

7. *There Is More Lawlessness and Violence*

Both land frontier and the nighttime have reputations as regions of danger and outlawry. Interestingly, both do not live up to the myths about them, for the patterns of aggression are selective and localized.

On the one hand there is clear evidence of lawlessness and violence. Walter P. Webb observed that the West was lawless "because the law that was applied there was not made for the conditions that existed. . . . It did not fit the needs of the country, and could not be obeyed" (cited by Frantz and Choate, 1955:83). There

was also a lack of policemen and law enforcement agencies were few (Riegel, 1947:627; Billington, 1949:480). There was violence in the gold fields (Hollon, 1974:211). In the cow towns, mining camps and boom towns in the early days, practically everyone carried guns. Fighting words, the ring of revolvers, and groans of pain were common sounds out there. Some western settlements were renowned for concentrations of gamblers and gougers and bandits, dance-hall girls and honky-tonks and bawdy houses. Horse thieving was widespread. The stage coach was held up many times. There was habitual fear of attack from either Indians or renegades. In the face of this, the people practiced constant watchfulness and banded together for self-protection (Billington, 1954:8; Doddridge, 1912:103). Towns had vigilante groups. The covered wagons that crossed the plains were accompanied by armed convoys.

Yet the violence was concentrated in certain places; otherwise killings and mob law were remarkably infrequent. Such infamous towns as Tombstone and Deadwood, and the states of Texas and California had more than their share of gunfights (Frantz and Choate, 1955:83; Billington, 1949:63; Hollon, 1973:96). But the tumult in the cow towns was seasonal, and took place when the cowboys finally reached Abilene, Ellsworth, and Dodge City after the long drive. And the mayhem was selective. Flint (1826:401) wrote, "Instances of murder, numerous and horrible in their circumstances, have occurred in my vicinity . . . in which the drunkenness, brutality, and violence were mutual . . . [Yet] quiet and sober men would be in no danger of being involved." W.T. Jackson (1973:79) adds, "Homicides and murders occurred so infrequently that when they did the community was shocked and outraged." Concerning violence, Hollon (1973:97-8) concludes that there was

a natural tendency to exaggerate the truth and emphasize the exception . . . not a single shoot-out took place on main street at Dodge City or any of the other Kansas cow towns in the manner of the face-to-face encounter presented thousands of times on television.

Why, then, did the land frontier have the

reputation of a "Wild West?" One reason may be that outlaw killers were drifters, so the same person may have contributed exploits over large areas. Another reason was boredom. The stories of violence persisted and spread because there was little to do or to read about in pioneer homes. The tedium of daily life was countered by exciting stories told and retold around the stove in the general store.

It is plausible that western desperados and nighttime muggers would have similar outlooks. Both believe there is less exposure, which improves their chances for succeeding at the risks they take. One relied on dry-gulching; the other uses the dark to set an ambush. Escape is easy because both could move from the scene of the crime into unpopulated areas and elude pursuers.

The nighttime has been noted also as a place of evil. It is thought of as crime-ridden and outside of ordinary social control. Medieval and Renaissance cities had no public illumination. Assaults by ruffians and thieves were so common after dark that wayfarers took to paying others to precede them through the streets carrying lighted torches. In the seventeenth century this escort-for-hire was called a "link boy" in London, and a "falot" (lantern companion) in Paris. Deliveries of black market goods to stores, such as fuel oil to gasoline stations during the oil embargo of 1973-1974, was accomplished under cover of darkness. Lawlessness is possible then because police coverage is sparse (Boston *Globe*, 1977:1). In addition, the officers on duty make themselves unavailable by sleeping in their cars, an old custom in New York City where the practice is called "cooping" (New York *Times*, 1968). The same was informally reported to me about Boston police as well; they are found snoozing in their police cars in the Arboretum by the early morning joggers.

In Boston today, carrying arms is more common at night. For fear of mugging or rape, escort services are provided on many college campuses for women returning to their dorms at night, or for women on the evening shift going from their places of work to the parking lot or subway station. An escort is provided for

nurses at Boston City Hospital because of an increase in robberies in that area. And some apartment houses, with their sentries at the door, become vertical stockades to which people in the city retreat at night.

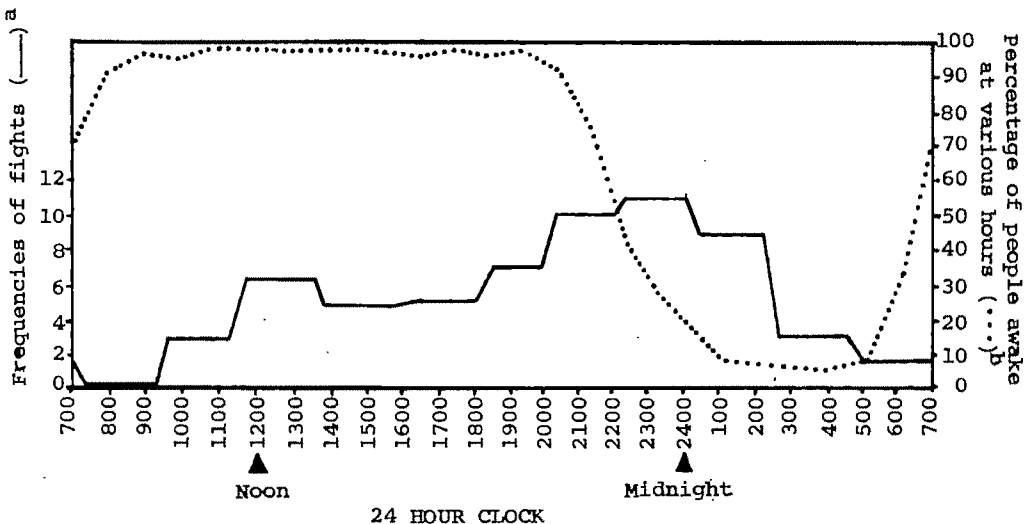
However, like the former West, lawlessness and violence at night are concentrated in certain hours in certain places and are otherwise uncommon. Fights reach their peak about midnight, as shown in Figure 3, but are least frequent from 2:30 to 11:00 a.m. The area of Boston in which many brawls and muggings take place, where prostitution is rampant and bars and lounges feature nude go-go dancers, is called the "combat zone." A large transient population of relatively young males come into the area to patronize the moviehouses featuring X-rated films and become drunk and aggressive in bars and on the streets. Although this description may approximate what was once reported of mining towns in the West, these combat zones do not function so after 2:30 a.m. or during the daytime. In the daytime the areas are parts of business districts. Many people shop at department stores nearby, or otherwise pass through and patronize eating places and businesses there. So the

combat zone designation refers to these places only at certain hours and is not true for all the city all night.

8. *There Is More Helpfulness and Friendliness*

Hollon (1974:211-2) remarks that "For every act of violence during the frontier period, there were thousands of examples of kindness, generosity, and sacrifice" He quotes an English traveler who said, "'Even the rough western men, the hardy sons of the Indian frontier, accustomed from boyhood to fighting for existence, were hospitable and generous to a degree hard to find in more civilized life.'"

Reports of life on the land frontier are replete with accounts of warmth toward strangers, of community house building and barn raisings, and of help for those in need (Darby, 1818:400; Frantz and Choate, 1955:64; Billington, 1949:96, 167; Riegel, 1947:81). "Neighbors were ready to lend anything they possessed. No man driving along with an empty wagon on a good road would pass another on foot without inviting him to ride" (Dick,



^a Fights reported in emergency calls to telephone operators in the city of Boston, based on a 49-day (24-hour daily) sample made in 1974 (June, September, and December). The frequencies are adjusted for sampling variations over the 24 hours.

^b Percentage of people awake is an estimate for Boston, and is based on a summary of 44 U.S. cities, reported in Szalai (1972:737).

Figure 3. Number of Fights Reported and Percentage of People Awake at Those Times in the City of Boston

1937:512). Travelers returning from the outskirts said they were treated more kindly than they had been in the cities (Flint, 1826:402-03; Hollin, 1974:212).

At first these stories of openhanded western hospitality may seem inconsistent in the face of the high risks of thievery and violence. But the circumstances are actually related to one another. Dick (1937:510) observed that "As the isolated settlers battled against savage men, . . . and loneliness, they were drawn together in a fellowship." Billington (1972:166) added,

Cooperation is normal within every in-group, but accentuates when the in-group is in conflict with an out-group and group solidarity is strengthened. This was the situation in frontier communities where conflicts with Indians, with raw nature, and with dominating Easterners heightened the spirit of interdependence.

That people want to affiliate under such conditions with others like themselves was demonstrated experimentally by Schachter (1959). He showed that the greater the risk people thought they were facing, the more anxious they were; and the more anxious they were, the more they wanted to be with others—even strangers—facing the same risk. Schachter (1959) concluded that being with others in the same boat served to reduce anxiety, and also provided an opportunity to appraise one's own feelings and adjust them appropriately to the risk. With less emotional uncertainty and with the knowledge that others share the circumstances, individuals feel better about confronting a stressful situation.

Because the night is a time of more violence and people feel more vulnerable then, those up and about have a similar outlook and behave toward others as pioneers did in the West. At night people are more alert to strangers when they pass on the street. Each tries to judge whether the other is potentially dangerous. Upon deciding that the other is to be trusted, one's mood shifts from vigilance to expansiveness. If not foe, then friend. Aware that they are out together in a dangerous environment, people identify with each other and become more outgoing. The sense of safety that spreads over

those together at night in a diner or in a coffee shop promotes camaraderie there.

Also, on both frontiers people may be more hospitable because they have time to devote to strangers. Pioneers had plenty to do; yet often they had nothing to do. They were not closely synchronized in daily tasks as people were in the eastern cities, and the norm of punctuality was not emphasized. One man who grew up in the West

. . . recalled the boredom he could never escape. . . . [T]he worst time of all was Sunday afternoon, when he had nothing to do. There were no newspapers to read and no books other than the family Bible, there was no one his age to talk with, and the nearest store was miles away. (Hollon, 1974: 196)

In the city during the day, the mood of pressured schedules takes hold of folk and makes their encounters specific and short. The tempo slows markedly after midnight. The few who are out then hurry less because there are fewer places to rush to. Whereas lack of time inhibits sociability and helpfulness, available time clears the way for them.

I checked on these ideas by four tests of people's helpfulness and friendliness at various times in the 24-hour cycle. The tests are modest situations, not emergencies to which one has to respond under stress, but part of the common stream of social events. The ratings for degree of helpfulness and friendliness were established by asking sets of individuals to act as judges (ten judges for Test 3, six each for Tests 1, 2, and 4).

Test 1: Asking for directions. A male and female couple used a random sampling procedure on the street² to approach

² Tests 1, 2, and 3 reported in this section were carried out during each of the 166 field visits for which the sample is described in Figure 2n. A team of two researchers, always one male and one female, made a field visit. The male always asked for directions, but the two took turns attempting to secure interviews (Test 2). Tests 1, 2, and 3 were carried out in locations that were illuminated after dark, for example under a street lamp. Tests 1 and 2 both called for random sampling of passersby, and the procedure used the five-minute person-tally that contributed the data reported in Figure 2. This tally ascertained the density of the street population at that time of day, and this density number was used to set the

passersby and ask directions to a well-known location about a mile away. This is a familiar question, calling for little time and effort and one does not have to become personal in replying. Giving directions was scored one point. A nasty reaction—such as brushing by stiffly as if ignoring a panhandler, or turning and uttering an obscenity or a terse “Ask someone else,” or quickly veering away from the speaker with apparent nonrecognition, or staring angrily and continuing to walk—was scored zero points. If the individual enlarged the scope of the encounter—such as by saying, in addition to giving directions, “Are you tourists here?” or “My son lives there,” or “Do you need a ride? I have my car nearby”—it was scored two points. After each trial the two field researchers came to an agreement about the rating. Of 363 persons approached, 331 (91.2%) gave only directions, 21 (5.8%) enlarged the encounter, and 11 (3%) refused nastily.

Test 2: Requesting a brief interview. A male and female couple followed a random selection procedure on the street (see fn. 2) to approach passersby and ask them to answer some questions in a survey being conducted about people who are in cities. This is a less common encounter but not unfamiliar in this land of frequent polls. One is asked to give time and trust to strangers, for the interview takes several minutes and the questions are personal—concerning feelings, employ-

ment, and living situation. A nasty refusal—such as an abrupt “Not from me [do you get an interview]!” or one of the reactions listed as nasty for Test 1 above—was scored zero points. A polite refusal—such as a plausible reason: “I’m sorry, I have an appointment” or equivalent delivered courteously—was scored one point. Consent was scored two points. Of 1,129 attempts, 175 (15.5%) refused nastily, 258 (22.9%) refused nicely, and 696 (61.6%) consented to the interview.

Test 3: Finding a lost key. A key found on the street is likely to be recognized as an object of value belonging to someone the finder does not know. Would the finder care enough about the stranger who lost it to send the key back? It is an anonymous situation. One does not have to become personal with the owner in order to return the key. This is a test of whether or not there is a difference in the rate of returning keys among people who pick them up at various hours and carry them away. The idea for using a key in such an experiment was developed by Forbes, TeVault, and Gromoll (1972). At the beginning of each of the field visits (see fn. 2), the researchers placed brightly colored aluminum keys in specific well-lighted locations on the streets at the sites. Each key had a tag attached listing the name and address of someone in a city ninety miles away (Northampton) and the request “Please return.” We avoided locations near store entrances since the keys might be turned in to a clerk rather than mailed directly. The keys were color-coded and notch-coded, so that we could link each returned key with a particular interval of the 24-hour period. At the end of each two-hour visit the researchers made the rounds and retrieved every key that had not been carried away.

Overall, of 326 keys carried away, 220 (67.5%) were sent back. If the key was dropped in the mailbox (from which they were delivered to Northampton and postage due charges paid), it was scored one point; 154 (47.2%) were returned in this manner. The 38 keys (11.7%) that were returned in a stamped wrapper were scored two points each. Another 25 stamped and wrapped keys (7.7%) came with a personal note enclosed and were

sampling rate: if the number was greater than twenty-five, the fourth person passing from the direction would be chosen once the test was to start. If the count was from five to twenty-five, the third person would be chosen; if three or four, then the second person would be chosen; and if two or fewer passed during the five-minute tally then the first person from either direction would be selected. The members of the same group passing by were counted as one person; if the group was sampled the passerby nearest the researchers was approached. A maximum of three requests for directions and eight requests for interviews was to be made during each field visit. Sometimes the attempts were fewer—especially at night—because almost no one passed by during the segment of the field-visit schedule in which these attempts were to be made. The sampling was silently carried out by one researcher, who then by nudging or an equally subtle signal told the other whom to approach. This minimized selection bias, for the researcher who was to carry out the test could not hesitate or overlook an unappealing passerby.

counted three points each. Three individuals (.9%) even telephoned Northampton to say that the key was in safe hands, and were scored four points each. The reactions reflect increased degrees of giving time and investing oneself in helping a stranger. The 106 keys (32.5%) not returned were scored zero points each.

Test 4: Being sociable in the supermarket. The supermarket checkout procedure is a microcosm of city street life. People who are mostly strangers to one another make limited contact in a brief, standardized situation. Research assistants, male and female couples, followed a stratified random sampling procedure³ to visit three 24-hour supermarkets in different parts of Boston at various hours of the day and night. They posed as customers and noted the degree of sociability between single customers and clerks at the checkout counters. Sociability is defined as "showing warmth and expanding the scope of interaction with another." In this case, smiling was taken as showing warmth. The criterion for expanding the scope of the encounter was chatting about a topic other than the transaction—such as saying to the clerk, "Are you a high school student?" or "That's a nice shirt; where'd you get it?" or reporting general news. If the customer both smiled and chatted, the encounter was scored two points. If either happened alone it was rated one point, and zero points were scored if neither took place. (Interaction between customer and clerk showed a high degree of mutuality. When one was impassive the other was too, and when

one was sociable the other responded in kind. That is why this simple scoring scheme was used rather than a tally that would have included the behavior of both parties.) To what extent do people break through the confines of their roles in the supermarket checkout procedure to offer even the mildest sociability? Not much, for 562 (74.7%) of the 752 customers did not smile or talk about something other than the transaction. Smiling by itself occurred 98 times (13%), and only 66 (8.8%) both chatted and smiled.

To summarize, over 2,500 people were observed in various parts of central Boston throughout the 24-hour cycle and were rated on how they responded to four situations: giving directions when asked, consenting to be interviewed when asked, returning lost keys they found, and being sociable with strangers during the focused moment of paying for goods at a supermarket checkout counter. Four tests were used so that several different behaviors would help define and give face validity to what is being studied. While these do not cover the entire range of helpfulness and friendliness, showing some warmth, cooperating with another's modest appeal, and expanding the scope of interaction are the initial conditions of such relationships.

The samples of people among the tests are not the same. Tests 1 and 2 used a periodic selection of passersby following a random procedure adjusted to street population density. Test 3 focused only on persons who carried keys away. Test 4 involves only single customers at the checkout register in always-open supermarkets. Nevertheless, direct *time* comparisons are appropriate, for the tests are all based on random sampling designs for the same intervals around the clock. The issue for evaluating the hypothesis will be the sizes of the differences found among times of day within each test and the consistency of results by time of day across the four tests.

The results of the tests are shown in Table 2. There is impressive consistency for three of the tests, with nighttime scores being highest. Not only does nighttime show up best in these three cases, there is no other time of day consistently

³ The supermarket sample was similar to the sample for the experiments on Boston streets (see Fig. 2n). The same time intervals were listed in the sample frame and thirty visits were selected randomly (three, morning rush; seven, daytime; four, evening rush; seven, evening; and nine, night). The quota for nighttime visits was higher because fewer customers would be observed on each visit then. The observers were eleven pairs of students from my course on Social Interaction in the Fall, 1973; they were trained and rehearsed via earlier trials to reliability averaging .92 (between pairs of observers). One supermarket is a quarter-mile from the residential site for the street experiments; the other two markets are located in different parts of town. In each visit observers followed a systematically-varied visiting order among the three stores, for a total of 90 fifteen-minute observation periods.

Table 2. Aggregate Findings of Four Separate Tests for Helpfulness and Friendliness during Various Phases of the 24-Hour Daily Cycle, in Boston in 1974

Test	Time of Day				Σ	Analysis of Variance and Significance
	Morning rush hour 0730-0929	Daytime 0930-1614	Evening rush hour 1615-1814	Evening 1815-0014	Night 0015-0729	
1. Give directions ^a	1.06 n=32 ^c	.97 n=115	1.07 n=27	1.00 n=123	[1.15] ^e n=66	F=4.917, p<.001
2. Consent to interview ^b	1.17 n=93	1.45 n=366	1.46 n=81	1.50 n=363	[1.55] n=226	F=4.531, p<.002
3. Return lost key ^c	.93 n=29	1.00 n=113	1.04 n=26	[1.18] n=94	.61(1) n=64	F=3.972, p<.01
4. Be sociable with stranger ^d	.24 n=68	.42 n=212	.29 n=161	.22 n=179	[.50] n=132	F=5.250, p<.001

^a Index based on zero points for nasty response and no directions given, one point for directions given only, two points for giving directions and expanding the scope of interaction as well; summed, and divided by number of trials (persons approached) within time period. Two of six judges differed in rating a courteous refusal—e.g., "I'm sorry, I don't know"—saying it is a common way to avoid getting involved. Hence polite refusals were omitted from the analysis. A check showed this made no difference. There were thirty instances of polite refusals, scattered over the times of day, and their inclusion (one point each) yielded the same ANOVA results as their exclusion.

^b Index based on zero points for nasty refusal, one point for polite refusal, two points for consent; summed, and divided by number of trials (persons approached) within time period.

^c Index based on zero points if key not returned, one point if returned unwrapped, two points if returned wrapped, three points if returned wrapped with message enclosed, four points if personal contact made by telephone; summed, and divided by number of trials (keys carried away) within time period.

^d Index based on one point for smiling only, one point for chatting only, two points for both, zero points for neither; summed, and divided by number of cases (transactions observed at the checkout counter) within time period.

* n = number of trials; [] = highest mean score for test.

second best. In some instances the differences between nighttime and its nearest competitor are not statistically significant, even though the analysis of variance yields significant results when all times are compared. Although differences among hours are small in given instances, the cumulative effect of these practices would make a noticeable difference in the social mood at various times. The overall pattern supports the prediction that nighttime is a period of more helpfulness and friendliness than other portions of the day.

In that light the outcome of the key test is surprising. The night had by far the lowest rate of helpfulness. The lowest proportion of keys were returned (50%) and the least extra effort, beyond dropping keys unwrapped into the mailbox, was made then. This finding is so clear-cut and contrary to expectations that it must be significant. Its interpretation would benefit from information still to be presented, and I will postpone comment about its bearing on the frontier hypothesis until later.

The pattern of findings for all four tests does reject a rival hypothesis: *fear* determines people's conduct toward strangers at night. We know the night is viewed as a dangerous time to be outside one's home in the city (U.S. Office of Management and the Budget, 1974:58-9, 73). If fear of criminal assault dominated social behavior then, it should be greater in face-to-face encounters than for the passive, anonymous appeal to find a key tagged "Please return." We would expect people to be more guarded towards others at night, to shun approaches by strangers, but to be more helpful in the low-risk situation of dropping a lost key into the mailbox. Table 2 tells us that just the opposite happened. Nighttimers were more helpful and friendly towards strangers face to face. And yet, of the keys picked up, they returned the fewest.

9. Exploitation of the Basic Resource Finally Becomes National Policy

Westward expansion began long before anyone officially recognized the land frontier's possibilities for our society. It took years to realize even that the U.S. West was habitable. At one time the land west

of the Missouri River was labeled on maps as the Great American Desert. Almost no one thought that some day many people would want to migrate and settle there (Hicks, 1948:508). Nor was the catch phrase "Manifest Destiny" applied to colonizing the West until 1845, centuries after the effort had been under way. In 1837 Horace Greeley introduced the slogan "Go West, Young Man, go forth into the Country." He looked upon such migration as a means of relief from the poverty and unemployment caused by the Panic of 1837. By 1854 Greeley was urging, "Make the Public Lands free in quarter-sections to Actual Settlers . . . and the earth's landless millions will no longer be orphans and mendicants" (cited in Smith, 1950:234-5). In 1862, with the passage of the Homestead Act, it became a deliberate policy of the U.S. government to use the western territory to help relieve the conditions of tenant farmers and hard-pressed city laborers. A member of Congress declared, in support of the Homestead Act, "I sustain this measure . . . because its benign operation will postpone for centuries, if it will not forever, all serious conflict between capital and labor in the older free states" (Smith, 1950:239). The policymakers finally saw the exploitation of western space as a means of solving social problems.

Similarly, in the first 150 years after Murdock's coal-gas illumination was introduced, there was no national consciousness in England or the United States about colonizing the nighttime. People went ahead, expanding their activities into the dark hours without declaring that a 24-hour community was being forged. Now in the 1970s policy makers have begun talking about cheap time at night the way they once spoke of cheap western land. V.D. Patrushev (1972:429) of the Soviet Union writes that "Time . . . is a particular form of national wealth. Therefore it is imperative to plan the most efficient use of it for all members of a society." Daniel Schydrowsky (1976:5), an economist who specializes in development in Latin America and who recently ended a three-year study there, has concluded that multiple-shift work would

produce remarkable gains in reducing unemployment and improve the economies of overpopulated developing cities. His claim for the use of time echoes the attitudes of nineteenth century proponents of the use of western lands as a solution for those who were out of work.

The advocates of westward expansion also saw it as a way to draw off great numbers of people from the cities and forestall crowding there (Smith, 1950:8, 238). Today Dantzig and Saaty (1973:190-3) recommend dispersing activities around the clock as a means of reducing congestion. And Meier (1976:965) writes, "Scarce land and expensive human time can also be conserved by encouraging round-the-clock operation. . . . By such means people can live densely without stepping on each other's toes."

10. Interest Groups Emerge

As the U.S. frontier matured, the population became more aware of its own circumstances and organized to promote its own concerns. Turner (1893:207; 1965: 54) remarked that the West felt a keen sense of difference from the East. He wrote:

. . . [F]rom the beginning East and West have shown a sectional attitude. The interior of the colonies was disrespectful of the coast, and the coast looked down upon the upland folk. . . . [The westerners finally] became self-conscious and even rebellious against the rule of the East. . . . [I]t resented the conception that it was merely an emanation from a rival North and South; that it was the dependency of one or another of the Eastern sections. . . . It took the attitude of a section itself. (1932:25-30)

Sections are geographically-based interest groups. One hundred years ago the West gave rise to such pressure groups and farm bloc organizations as the Greenback Party, the National Grange, and the Populists. The Granger movement, for example, grew with the westerners' problems with transportation in their region. There were no significant river or canal systems out west and so the settlers were at the mercy of railroads. But the rates in the newer regions of the West were far higher than those in the

East, and it was protest against this disparity that aided the movement in the 1870s (Robbins, 1960:271).

The night also isolates a group from the main society. Antagonism may develop as daytimers deprecate the nighttimers and the latter resent the neglect shown by the others. People active after dark find their life style differing from that of daytime society, become aware of having a separate identity, and evolve into interest groups. New alignments in the tradition of sectionalism begin to emerge. This has already happened for two groups usually linked with the nighttime: homosexuals and prostitutes. The Gay Liberation Front is one nationwide organization devoted to the rights of homosexuals. Prostitutes also have a union. Appropriately they adopted the name of a creature renowned in the U.S. West for howling at night—the coyote. COYOTES (Call Off Your Old Tired Ethics) seek legislation to decriminalize their activities and protest courtroom discrimination against women who earn their living by prostitution (Boston *Globe*, 1976a).

An actual day vs. night contest has already been fought in Boston. The city's airport is flanked by residential neighborhoods and its afterdark activity became a nuisance to people wanting an undisturbed night's sleep. In 1976 dwellers in those neighborhoods, as private citizens and through two organized groups—Fair Share, and the Massachusetts Air Pollution and Noise Abatement Committee—made a concerted effort to stop airplane flights between 11 p.m. and 7 a.m. It led to counterarguments by the business community stressing the economic benefit of continuing the flights. The pro-nighttime group was a coalition among commercial interests, airline companies, unions, and airport employees holding jobs at night (some of whom lived in those very neighborhoods). This group argued that the curfew would result in the loss of thousands of jobs, millions of dollars in sales, and further would discourage business investment in the New England area. Joined by the governor, the mayor and many legislators, the coalition successfully won a decision from the Massachusetts Port Authority that the nighttime flights should

be kept going. (Some proposals for noise reduction during the night accompanied the decision.) A month later, Eastern Airlines announced it was adding an airbus and expanding its staff at the airport "as a direct result of the recent decision . . . not to impose a night curfew at Logan [airport]." As one businessman put it, "The curfew decision was regarded as the shootout at the OK Corral" (Boston *Globe*, 1976b; 1976c).

DISCUSSION

The evidence bears out the hypothesis that night is a frontier. That nighttimers are *less* likely to return the keys they find also supports the idea. While the outcome of Test 3 seems to deny the claim that more help is given on a frontier, the lost-key experiment differs from the other tests in that it is the only one in which people do not meet face to face. It is a test of anonymous helpfulness. During the nighttime, strangers identify more readily with one another. A young man told me, "At 4 a.m. if someone sees you walking the streets at the same time he does, he must think, 'Gee, this guy must be part of the brethern, because no one else is awake at these times.' " However, if someone finds a key and does not know the owner, he would guess that everyone who passed that way is equally likely to have lost it. Nighttimers, knowing they are few, assume on the weight of numbers that the person who lost the key is a daytimer. In item ten above, I suggested that the feelings of nighttimers toward daytimers resembled the attitudes of westerners toward easterners a century ago. They perceive they are different and resent the neglect shown by the day people toward them. The nighttime in-group feels comradesly within itself but indifferent or antagonistic toward the out-group (see Sumner, 1906:27). Whereas frontier people readily help others whom they meet on the frontier, their sense of difference from unknown daytimers leaves them less concerned about the others' plights and they do not return many lost keys.

I cannot think of an equally plausible rival explanation, compatible with the rest

of the evidence, for this finding. This interpretation makes sense of the complete set of outcomes in Table 2 and fits the analysis in the preceding section. Revealing patterns stand out. One is the connection between violence and helpfulness and friendliness, a condition that emerges on the frontier because of fear there and solidarity among those who believe they share the dangers together. Another is the pairing of sectional attitudes and helpfulness, so that assistance is given selectively to those with whom the individuals identify.

The experiments confirm what we know about life on frontiers, but I did not explore wholly the causes of behavior here. The findings may be compared with research on helpfulness reported by Bryan and Test (1967), Feldman (1968), Latané and Darley (1970), Milgram (1970), Wispé and Freshley (1971), Darley and Batson (1973) and others. There is a problem of comparability because different times of day were not treated systematically in those studies. Yet some of the insights may work well together. The findings about available time, at least, agree with each other. Darley and Batson varied the degree to which their subjects were hurrying to an appointment when they came upon a person coughing, groaning, and apparently needing help. Of several possible influences that were measured, including what was in the subjects' thoughts at the moment (some of them were preparing to discuss the Good Samaritan parable!), only the degree of hurry was related to helping. A mere 10% of those who were late to their appointments stopped to help, whereas 63% of those who had ample time stopped to give aid to the crouching and suffering man.

CONCLUSION

What is the gain in thinking of night as a frontier? A single theoretical idea gives coherence to a wide range of events: the kind of people up and about at those hours, why they differ from daytimers in their behavior, the beginnings of political efforts by night people, the slow realization among leaders that public policy might be applied to the time resource.

Even the variety of endeavors becomes understandable—from metal smelting plants to miniature golf courses, to mayor's complaint offices, to eating places, to computerized banking terminals that dispense cash. The niche is being expanded. Bit by bit, all of society migrates there. To treat this as a sequel to the geographic spread of past centuries is to summarize the move within familiar ecological concepts of migration, settlement, and frontier.

Though I have reviewed materials for one period in U.S. history, these conditions are features of all frontiers. They should apply to the Russians crossing the Urals, to the Chinese entering Manchuria during the Ch'ing dynasty, to the Boers settling South Africa, to Australians venturing into the Outback, to present-day Brazilians colonizing the Amazon interior, as well as to Americans migrating into the night. The patterns are confirmed by essays in Wyman and Kroeber's anthology on frontiers.

We should also consider the uniqueness of this new frontier. Each settlement beyond established boundaries has its own qualities. Here are some differences between the West and the night: (1) On the land frontier settlers lived rudely with few services at hand. At night a large portion of the total range of activities is services. (2) Utilities cost more on the western fringes; at night the fees for telephone calls, electricity, and airplane travel are lower. (3) While western settlements were in remote contact with the East, day and night are joined so that either can be affected quickly by events in the other. Twenty-four hour society is more constantly adjusting, more unstable. (4) Looking westward, pioneers saw no end to the possibilities for growth, but we know that expansion into the night can only go as far as the dawn. (5) The land frontier held promise of unlimited opportunity for individuals who ventured there. Miners and pioneers endured hardships because they lived for the future. They hoped to make their fortunes, or at least a better life. At night there are large numbers of unskilled, menial, and dirty tasks; but charwoman and watchman and hospital aide and porter are dead-end jobs. Many people so

employed are immigrants or members of minority groups and this expanding margin of society is a *time ghetto*. The ghetto encloses more than minorities and immigrants, for ultimate control in 24-hour organizations remains with top management in the daytime. Policy making, important decisions, employee hiring, and planning are curtailed during off-hours. Since evening and night staffs are prevented from taking many actions that would lead to the recognition of executive ability, and since their performance is not readily observable by the bosses, all have poorer chances for advancement. (6) The western frontier's natural resources were so extensive that we became wasteful and squandered them. At night there is nothing new to exploit but time itself, so we maximize the use of fixed assets and become more frugal. (7) Migrating westward called for rather significant capital investment—outlays for a covered wagon, mining equipment, cattle, the railroad. There is little extra capital required for a move to the night. Instead, the incessant organization's need for more personnel reflects a swing toward more labor intensive operations. So the night frontier may appeal to developing countries with meager treasuries and teeming populations of unemployed.

This expansion is also unusual because it happens in time rather than in space. We change from a diurnal into an incessant species. We move beyond the environmental cycle—alternating day and night—in which our biological and social life evolved, and thus force novelty on these areas. (8) In the past a single set of minds shut down an enterprise one day and started it up the next. It permitted easy continuity and orderly administration. For coverage around the clock, we introduce shifts of personnel. Several times a day another set of minds takes over the same activity and facilities. (9) A physiological upset is imposed on people who work at night and maintain ordinary recreation and social life on their days off. Each time they switch their active hours they undergo phase shifts in body rhythms such as heartbeat, temperature, and hormonal production. The several days' malaise that results was known to such

workers long before air travel across time zones popularized the phrase "jet fatigue."

Ibsen's (1890: Act II) character, Eilert Lövborg, describes the two sections of the book he has written, "The first deals with the . . . forces of the future. And here is the second forecasting the probable line of development." We may believe we understand the forces, the conditions under which humans enlarge their niche, but what is the probable line of development? Forecasting is called for despite the difficulties of social prediction. We should consider the possibilities of an era in which unremitting activity is even more commonplace. What is the carrying capacity of the 24-hour day? What will happen when saturation occurs? Time will have extraordinary leverage as it gets used up, for time is a resource without direct substitute. It is unstretchable; we cannot do with it as we did with land by building up toward the sky and digging into the ground. Time is unstorable; we cannot save the unused hours every night for future need.

In his essay "The Frontier in American History," Frederick Jackson Turner (1893:38) reviewed the impact of the advance into western lands upon our society and remarked, "And now, four centuries from the discovery of America, at the end of a hundred years of life under the constitution, the frontier has gone." But it has not gone. During the era that the settlement of our land frontier was being completed, there began—into the night—a large-scale migration of wakeful activity that continues to spread over the world.

REFERENCES

- American Heritage Dictionary of the English Language
1970 Boston: Houghton Mifflin.
- Becker, Howard
1963 *Outsiders: Studies in the Sociology of Deviance*. New York: Free Press.
- Billington, Ray Allen
1949 *Westward Expansion*. New York: Macmillan.
1954 *The American Frontiersman*. London: Oxford University Press.
1972 "Frontier democracy: social aspects." Pp. 160-84 in G. R. Taylor (ed.), *The Turner Thesis: Concerning the Role of the Frontier in American History*. 3rd ed. Lexington, Ma.: Heath.
- Boston Globe
1976a "Prostitutes speak of pride, but they are still victims." June 25:1, 10.
1976b "Dukakis decides to go against Logan curfew." August 12:1, 20.
1976c "Logan anti-noise plan offered." August 13:35.
1977 "Boston police today." April 4:1, 3.
- Brandeis, Louis D. and Josephine Goldmark
1918 *The Case Against Night Work for Women*. Rev. ed. New York: National Consumers League.
- Bryan, James H. and Mary Ann Test
1967 "Models and helping: naturalistic studies in aiding behavior." *Journal of Personality and Social Psychology* 6:400-7.
- Buttimer, Anne
1976 "Grasping the dynamism of lifeworld." *Annals of the Association of American Geographers* 66:277-92.
- Dantzig, George B. and Thomas L. Saaty
1973 *Compact city*. San Francisco: Freeman.
- Darby, William
[1818] "Primitivism in the lower Mississippi valley." Pp. 399-401 in M. Ridge and R. A. Billington (eds.), *America's Frontier Story*. New York: Holt.
- Darley, John and C. Daniel Batson
1973 "From Jerusalem to Jericho: a study of situational and dispositional variables in helping behavior." *Journal of Personality and Social Psychology* 27:100-8.
- Dick, Everett
[1937] *The Sod-House Frontier, 1854-1890*. New York: Appleton-Century.
- Doddridge, Joseph
[1912] "Life in the old west." Pp. 101-6 in M. Ridge and R. A. Billington (eds.), *America's Frontier Story*. New York: Holt.
- Feldman, Roy E.
1968 "Response to compatriot and foreigner who seek assistance." *Journal of Personality and Social Psychology* 10: 202-14.
- Flint, Timothy
[1826] "Frontier society in the Mississippi valley." Pp. 401-3 in M. Ridge and R. A. Billington (eds.), *America's Frontier Story*. New York: Holt.
- Forbes, Gordon. B., R. D. TeVault, and H. F. Gromoll
1972 "Regional differences in willingness to help strangers: a field experiment with a new unobtrusive measure." *Social Science Research* 1:415-9.
- Frantz, J. B. and J. E. Choate
1955 *The American Cowboy: The Myth and Reality*. Norman: University of Oklahoma Press.
- Godkin, Edwin L.
[1896] "The frontier and the national character." Pp. 13-6 in M. Ridge and R. A. Billington (eds.), *America's Frontier Story*. New York: Holt.

- Hicks, John D.
1948 *The Federal Union*. Boston: Houghton Mifflin.
- Hollon, W. Eugene
1973 "Frontier violence: another look." Pp. 86-100 in R. A. Billington (ed.), *People of the Plains and Mountains*. Westport, Ct.: Greenwood Press.
1974 *Frontier Violence*. New York: Oxford University Press.
- Humphrey, Seth K.
1931 *Following the Prairie Frontier*. Minneapolis: University of Minnesota Press.
- Ibsen, Henrik
[1890] Hedda Gabler. Tr. E. Gosse and W. Archer. Pp. 42-74 in J. Gassner (ed.), *A Treasury of the Theatre*. New York: Simon and Schuster.
- Jackson, W. Turrentine
1973 "Pioneer life on the plains and in the mines." Pp. 63-85 in R. A. Billington (ed.), *People of the Plains and Mountains*. Westport, Ct.: Greenwood Press.
- Kozak, Lola Jean
1974 "Night people: a study of the social experiences of night workers." *Michigan State University, Summation* 4:40-61.
- Latané, Bibb and John Darley
1970 *The Unresponsive Bystander*. New York: Appleton-Century.
- Marx, Karl
[1867] *Capital*. New York: Modern Library.
1906
- Meier, Richard L.
1976 "A stable urban ecosystem." *Science* 192:962-8.
- Melbin, Murray
1977 "The colonization of time." In T. Carlstein, D. Parkes, and N. Thrift (eds.), *Timing Space and Spacing Time in Social Organization*. London: Arnold.
- Milgram, Stanley
1970 "The experience of living in cities." *Science* 167:1461-8.
- Mumford, Lewis
1961 *The City in History*. New York: Harcourt Brace.
- National Industrial Conference Board
1927 *Night Work in Industry*. New York: National Industrial Conference Board.
- New York Times
1968 "'Cooping': an old custom under fire." *December 15: Sec. 4, 6E*.
- Patrushev, V. D.
1972 "Aggregate time-balances and their meaning for socio-economic planning." Pp. 429-40 in A. Szalai (ed.), *The Use of Time*. The Hague: Mouton.
- Riegel, Robert E.
1947 *America Moves West*. New York: Holt.
- Robbins, Roy M.
1942 *Our Landed Heritage*. Princeton: Princeton University Press.
- Schachter, Stanley
1959 *The Psychology of Affiliation*. Stanford: Stanford University Press.
- Schlesinger, Arthur
1933 *The Rise of the City: 1878-1895*. New York: Macmillan.
- Schydrowsky, Daniel
1976 "Multiple shifts would produce 'revolutionary results' for Latin American economy." *Boston University, Spectrum* 4 (September 9):5.
- Smith, Henry Nash
[1950] *Virgin Land*. New York: Vintage.
1957
- Sumner, William Graham
[1906] *Folkways*. New York: New American Library.
1960
- Szalai, Alexander (ed.)
1972 *The Use of Time*. The Hague: Mouton.
- Turner, Frederick Jackson
[1893] *The Frontier in American History*. New York: Holt.
1920
1932 *The Significance of Sections in American History*. New York: Holt.
[1965] *America's Great Frontiers and Sections*.
1969 *Unpublished essays* edited by W. R. Jacobs. Lincoln: Nebraska University Press.
- U.S. Bureau of the Census
1975 *Historical Statistics of the United States, Vol. 1: ser. A195-209*. Washington, D.C.: U.S. Government Printing Office.
- U.S. Bureau of Labor Statistics
1976 *Current Population Survey*. Unpublished paper. May 12: Table 1. Washington, D.C.
- U.S. Office of Management and Budget
1974 *Social Indicators, 1973*. Washington, D.C.: U.S. Government Printing Office.
- Wispé, Lauren G. and Harold B. Freshley
1971 "Race, sex, and sympathetic helping behavior: the broken bag caper." *Journal of Personality and Social Psychology* 17:59-64.
- Wyman, Walker D. and Clifton B. Kroeber (eds.)
1957 *The Frontier in Perspective*. Madison: University of Wisconsin Press.
- Young, Michael and Peter Willmott
[1973] *The Symmetrical Family*. Harmondsworth, England: Penguin.

THE DIFFUSION OF COLLECTIVE VIOLENCE*

BRIAN L. PITCHER, ROBERT L. HAMBLIN AND JERRY L. L. MILLER

University of Arizona

American Sociological Review 1978, Vol. 43 (February):23-35

Past explanations of violence have characteristically paid more attention to the issues of social conditions and psychocultural stimuli than to the issue of timing. Timing is the focus of this paper in which a differential equation model of the temporal diffusion of violence is developed. This model is derived from behavioral generalizations which indicate that aggression is both instigated and inhibited via direct and vicarious learning. The parameters of the model provide measures of the instigation and inhibition processes that take place throughout an outbreak. Twenty-five data sets representing a wide variety of collective outbreaks of violence are used to test the empirical fit and to evaluate the credibility of the assumptions of the model. The model describes the overtime distribution of incidents quite accurately and the assumptions and implications of the derivation appear to be consistent with the cultural conditions surrounding the outbreaks.

One basic shortcoming of research on causes of violence is the assumption of independence among incidents. This assumption overlooks the fact that present actions are to a large degree affected by the outcomes of actions experienced in the past, either directly or vicariously. Explanations that simply identify different social conditions which cause violence, such as social disorganization (Downes, 1968), absolute deprivation (Lupsha, 1969), political structure (Lieberson and Silverman, 1965), organizational capacity (Shorter and Tilly, 1974; Snyder, 1975) and social conflict (Sears and Tomlinson, 1968) or that specify different psychocultural stimuli of violence such as relative deprivation (Gurr, 1968; Caplan and Paige, 1968), rising expectations (Davies, 1969), alienation-powerlessness (Kerner, 1968), internal-external control (Gurr, 1970) and

normative or value conflicts (Spiegel, 1971) are incomplete. They fail to take into account the time related social learning processes that mediate the impact of these factors.

A number of researchers have analyzed time-series data to investigate cyclical outbreaks of various kinds of violence (c.f. Denton and Phillips, 1968; Huff and Lutz, 1974; Li and Thompson, 1975; Lieberson and Silverman, 1965; Shorter and Tilly, 1974; Snyder, 1975; Snyder and Kelly, 1976; Spilerman, 1970; 1971; Tilly et al., 1975). Some have discussed possible interdependency or contagion effects, and a few (Huff and Lutz, 1974; Li and Thompson, 1975; Midlarsky, 1970) have attempted to measure and evaluate contagion processes. Spilerman (1970; 1971), for example, produced strong evidence that the probability of contagion in the U.S. urban riots was proportional to the population size of relevant units. However, social contagion has not been rigorously conceptualized in terms of social learning processes where individuals are instigated and inhibited by the outcomes of others' violent actions.

The purposes of this paper are to develop a differential equation model of the diffusion process inherent in most outbreaks of violence and to evaluate the adequacy of the model and its implications. The adequacy will be judged by: (1) empirically fitting it to data on a variety of

* This is a revision of a paper presented at the International Workshop on Comparative Ecological Analysis of Social Change, Ljubljana, Yugoslavia, August, 1976. The research was supported by Grant Soc 75-08448 from the National Science Foundation. The specification of the nature and functional form of the model was mainly done by the second author. The authors wish to thank their colleague, Michael Hout, for reading the manuscript and making a number of helpful criticisms and suggestions. An in-house reviewer gave an earlier version of the manuscript a careful reading and produced three pages of queries which helped us considerably in the revision. We are very appreciative of that effort and expertise.

outbreaks of violence; (2) comparing the fit of alternative differential equation models derived using different inhibition terms; (3) examining the congruence between the model's assumptions and known facts about particular outbreaks; and (4) discussing the reasonability and meaning of the parameter estimates.

THEORY

Violence may be defined as any activity which results in the nonaccidental physical damage of persons or property. It generally, if not always, occurs in a conflict situation where two or more parties contest to settle an issue in their own favor. There are several kinds of contests (e.g., legal and political, as well as violent). In a violent conflict, the parties damage each other's persons and property until one is destroyed, concedes the issue, or a compromise or stalemate is reached. Some violence is individual as when a husband beats his wife. Other violence may be organized as when two or more armies battle each other. Collective violence is a mixture of the two—unorganized individuals, collectivities or organizations involved in a battle over time against a common foe (the country, the establishment, the government, adults, blacks, or Jews, for example). It involves social contagion wherein the units are instigated and inhibited by the information they receive through time about one another's behavior and its consequences. Hence, the modifier collective here implies a population of units separated in time and space and influencing one another as they act together on the basis of secondhand information and without hierarchical leadership.

Collective Violence as Cultural Diffusion

It has been suggested (Huff and Lutz, 1974) that the logistic model of cultural diffusion (Dodd, 1953; 1955; Griliches, 1957; Coleman et al., 1966) explains why violent incidents usually accumulate in a sigmoid pattern. However, a careful

analysis reveals four serious problems with this formulation.

First, the mechanism in the logistic model is communication by salesmen and admen, and people who have already adopted and experienced the benefits and costs (cf. Coleman et al., 1966). In contrast, the probable mechanism for outbreaks of collective violence is much more subtle: imitation or vicarious learning usually based on news media reports of violence and its consequences occurring in other parts of the nation or world (Archer and Gartner, 1976; Spilerman, 1976).

Second, while the cumulative curves of collective violence are sigmoid, many violate an essential condition of Dodd's (1953; 1955) and Coleman et al.'s (1966) logistic theory: the units of the populations which generate them are seldom, if ever, in *direct* communication with one another. Consequently, the redundancy of contacts cannot explain the later slowing of the accumulation that results in the sigmoid pattern. According to the Coleman et al. (1966) theory, without direct communication among units in the population, the data should take a decaying exponential shape, not the sigmoid pattern which all exhibit.

Third, the logistic diffusion distribution results when each unit's adoption or first use of the invention is counted, whereas the sigmoid curves of collective violence are generated by actors who often participate more than once. For example, in the sigmoid outbreak of coups d'état among countries in Africa, 1960–1975, one of the units (Dahomy) produced at least one coup during each of seven years.

Fourth, while the violence data are more or less sigmoid, they are not all symmetric. This asymmetry excludes the logistic model from serious consideration since it is always symmetric.¹ The suggestion by

¹ The logistic is symmetrical because it is assumed that the rate of adoption is a constant. If, alternatively, it is assumed that this rate changes through time, the model becomes asymmetrical (Hernes, 1972). This necessitates the introduction of an additional parameter. Generally the asymmetric Logistic is not preferred over more parsimonious models, such as the Gompertz, unless the fit to the data is consistently and significantly better.

Hamblin et al. (1973) that a logistic model of innovation describes outbreaks of violence may be rejected for the same reason.

For the above reasons, it appears that the nature and mechanisms of the diffusion of collective violence are not isomorphic with those of general cultural diffusion, and that an alternative model is needed.

Mechanisms of Collective Violence Diffusion

There have been a number of suggestions (cf. Turner and Killian, 1972) regarding specific mechanisms for the spread or diffusion of collective behavior: suggestibility, circular reaction, identification and, as noted, imitation. While all of these may have some merit, the imitation mechanism has enjoyed by far the most theoretical and empirical investigation, with most of the recent developments pioneered by Albert Bandura (cf. 1977) and his colleagues.

Imitation or modeling involves vicarious learning. By watching others working the environment or by talking with them about their experiences, people are made aware of reinforcing consequences, and thereby learn what works and what does not work. However, according to Bandura's (1973) experimental results, learning is not to be confused with behaving. A subject may have vicariously learned that a particular behavior produces a specific reinforcement. Yet, to engage regularly in that behavior the subject ordinarily must frequently encounter the same cues, observe models who legitimate the behavior and personally try the behavior and experience the reinforcing consequences. Evidence from several experiments suggests that overt aggression occurs with substantial frequency only when people are threatened in a conflict situation and observe a model successfully aggressing against the source of threat, the other party in the conflict. The threat without the aggressing model or the aggressing model without the threat produces minimal aggression (Wheeler and Caggiula, 1966; Hanratty et al., 1972). Other experimental data (cf. Bandura et al., 1963; Gilmore, 1971) show that in a threatening conflict the observers' aggressive be-

havior increases as the success of the model's aggression increases (where success is defined as resolving the issue in favor of the model and the model's getting by without punishment).

The Mathematical Derivation

In large scale outbreaks of collective violence much observational learning evidently occurs secondhand, via reports and descriptions in the mass media. By attentively following news media accounts of incidents of violence, units affected by the conflict become familiar with the methods and innovations used by both the units who have thus far participated and by the combatants. The following derivation assumes that a more or less constant proportion of the violent events occurring during an outbreak are reported in the media (see Snyder and Kelly, 1977).

The derivation starts with a definition equating three basic concepts relating to the imitation mechanism. A party in a conflict is composed of a population of behaving units, individuals, groups, collectivities or organizations. Given the assumption that the timing of a unit's participation in collective violence is determined primarily by observational and symbolic learning vis-à-vis others, each violent incident by a unit is both an imitation of previous behaviors and a behavioral model for other units to imitate.² Therefore, at any time (t), the cumulative number of violent events (V) in the outbreak equals the cumulative number of imitations (I) by units and the cumulative number of behavioral models (M) for units, or:

$$V = I = M. \quad (1)$$

The social psychologists involved in the experimental work on imitation have not attempted to develop equations. However, sociologists (cf. Dodd, 1953; 1955; Coleman et al., 1966) in deriving the logistic model of cultural diffusion have

² It is assumed that the first unit to start the outbreak is imitating another unit outside the system or, much less frequently, has independently invented the particular kind of violence in question.

used an exponential differential equation which, given epistemic assumptions similar to the above, is an equation for imitation. This exponential differential equation for imitation specifies that the increment in imitations (dI) per increment in time (dt) is some proportion (p) of the number of previous imitations (I):

$$\frac{dI}{dt} = p I \text{ or } \frac{dI}{I} = p dt. \quad (2)$$

Because of the equalities in (1), equation (2) implies:

$$\frac{dM}{M} = p dt \quad (3)$$

and

$$\frac{dV}{V} = p dt. \quad (4)$$

The parameter (p) is the rate at which imitation is instigated and is assumed to depend on the costs and benefits of that particular action relative to those of all plausible alternative actions. This is consistent with both experimental evidence on the imitation of aggression (Bandura, 1973) and evidence for the relationship between the rate of diffusion and amount of reinforcement (Hamblin and Miller, 1976).

Equation (4) involving the collective violence terms (dV and V) is the one of principal interest because of the close correspondence between these terms and the data. However, the others are important because they specify the vicarious learning and imitation processes from which (4) is derived.

Aggression is not only instigated through behavioral modeling but may also be inhibited by costs resulting from failures and the deterrence strategies implemented by opposition units (cf. Bandura, 1973). Increases in the frequency and magnitude of such costs relative to alternatives augment the number of instigated units who are inhibited from engaging in the violence. Inhibition effects are assumed to be cumulative so the relative increase in violence (dV/V) per increment of time (dt) not only varies directly with the instigation rate (p) but also inversely

with the accumulated number who are inhibited (i) as in the following equation:

$$\frac{dV}{V} = \frac{p}{i} dt. \quad (5)$$

In outbreaks of collective violence relative costs are observed and talked about and the experimental data show that aggressive behavior by threatened observers is inhibited as they observe the punishment of aggressing models (Bandura et al., 1963; Gilmore, 1971). It is, therefore, assumed that observational and symbolic learning occurs to inhibit violence and that the imitation equation (2) applies so that the increment in numbers inhibited per increment of time (di/dt) is some proportion (q) of the accumulated number of units who have been inhibited (i) up to time (t), or:

$$\frac{di}{dt} = q i, \frac{di}{i} = q dt. \quad (6)$$

Solving (6) for i via integration yields:

$$i = i_0 e^{qt}, \quad (7)$$

where q is the previously defined rate at which units are inhibited, e is the base of the natural logarithm and i_0 is the value of i when $t = 0$. This value of i may be substituted into (5) to obtain:

$$\frac{dV}{V} = \frac{p dt}{i_0 e^{qt}} = c e^{-qt} dt$$

or

$$\frac{dV}{dt} = c e^{-qt} V \quad (8)$$

where $c = p/i_0$, or the net rate at which units are instigated to imitate the violence in question, and q is the rate at which they are inhibited. Thus, if this model adequately describes the through-time diffusion of a particular type of violence, the nonlinear regression analysis will provide measures of the rates of the two constitutive processes—instigation and inhibition—for the 25 data sets modeled here.

THE ANALYSIS

The Data

In order to evaluate the descriptive adequacy of the model, we obtained data

sets on ten different forms of collective violence: lynchings of blacks, vandalism and swastika painting on Jewish buildings, air hijacking attempts and attempts to deter hijackings, guerrilla warfare, revolutions, purges, coups d'etat, agrarian protests and civil disorders.³ For a description of the data see Table 1.

The Estimation of Parameters

Because both instigation and inhibition are assumed to approximate continuous rather than discrete processes, the model is stated in differential rather than difference equations. The term (dV/dt) refers to the rate of violence during very small increments of time and V is the accumulated number of incidents of collective violence up to a point in time. The accumulated data correspond to V very much better than data per day, week, etc. correspond to dV/dt , and for that reason, equation (8) was solved by integration for V :

$$V = V_0 e^{(c/q)} e^{(-c/q)e^{-qt}} = A e^{kb^t} \quad (9)$$

The right part of the equation was used in fitting the data and c and q were calculated using the following identities: $c = -qk$ and $q = -\ln b$, which are implied above. The intermediate steps required to obtain

(9), usually referred to as the Gompertz equation, were omitted because they involve calculus and are quite technical. Those proficient in calculus can derive the equation for themselves; those not would scarcely be helped by the rationale that could be outlined in the space available here. The essential point is that by fitting the integrated equation to the accumulated data, one is able to calculate estimates of c and q as postulated in the model.

The fitting or estimating was done with a nonlinear least-squares regression program based on the Fletcher-Powell (1963) optimization technique. Equation (9) was also fit to most of the data sets using the SPSS nonlinear regression program, with identical results to the fourth significant digit.

RESULTS AND DISCUSSION

Goodness of Fit

The results of the nonlinear least-squares regression analyses are given in Table 1. The integrated equation fits the data very well: r^2 values range from .941 to .999 with a median of .995. Overall, there are no systematic deviations in the residuals. The empirical fit to the accumulated distribution of events is quite acceptable (i.e., $r^2 > .98$) for all but the 1967 U.S. outbreak of civil disorders.

Fit of Alternative Differential Equation Models

Two alternative sigmoid models were considered but rejected. The alternative models were:

$$\frac{dV}{dt} = sVt^m \longrightarrow$$

$$V = \left(\frac{V_0}{e^{s/(m+1)}} \right) e^{(s/(m+1))t^{m+1}} \quad (10)$$

$$\text{and } \frac{dV}{dt} = rV(N - V) \longrightarrow$$

$$V = \frac{N}{1 + \frac{N - V_0}{V_0} e^{-Nrt}} \quad (11)$$

Both of these were derived using instiga-

³ All apparent indications are that the various coders took considerable care to assure the accuracy of each of these data sets. For example, Banks (1971) reports that the intercoder stability averaged .974. However, independently coded data are available for the air hijacking attempt outbreaks and the 1958-66 outbreak of coups d'etat in Latin America. The agreement between the FAA reports on hijacking and the reports in the *New York Times* on worldwide hijacking was over 98 percent. In addition to Banks, data on successful Latin American coups d'etat are also available in Li and Thompson (1975) and Solaún and Quinn (1973). The only discrepancy among all three data sets concerned the coding of the Cuban revolution-coup (Banks coded it in 1959, Solaún and Quinn coded it in 1958 and Li and Thompson did not include it). In addition, Li and Thompson coded two events as coups that the others did not—one in Brazil, 1961, where the President resigned under severe pressure and a second in Peru, 1963, where the head of the military junta was replaced by another General. The high overall agreement among data sets is evidence of their accuracy. Also, in modeling, measurement error typically attenuates the r^2 . Therefore, the consistently high r^2 values here evidence the data as both reliable and valid.

Table 1. Nonlinear Least-Squares Parameter Estimates and R² Values for 25 Collective Violence Outbreak Data Sets

Collective Violence Outbreak	Duration	Equation: $V = V_0 e^{(c/q)} e^{(d/-q)e^{-qt}}$				Data Sources
		$V_0 e^{(c/q)*}$	c	q	r ²	
Lynching of blacks, U.S. ^a	1882-1956	3483.00	.0007	.0002	.998	U.S. Bureau of the Census, (1960)
Anti-Semitic, U.S., 1958 ^b	10/12/58-11/29/58	84.29	.2985	.1395	.982	Caplovitz and Rogers (1961)
Anti-Semitic, U.S., 1960 ^b	12/26/59-2/29/60	643.30	3145	.1046	.995	Caplovitz and Rogers (1961)
Air hijacking attempts, U.S. ^c	11/67-12/69	73.01	.0622	.0096	.994	Federal Aviation Agency (1974)
Air hijacking attempts, U.S. ^c	1/70-8/71	53.49	.0232	.0080	.997	Federal Aviation Agency (1974)
Air hijacking attempts, U.S. ^c	7/71-7/73	53.29	.0531	.0148	.995	Federal Aviation Agency (1974)
Air hijacking attempts, U.S. ^c	9/67-10/70	104.20	.0332	.0049	.995	Federal Aviation Agency (1974)
Air hijacking attempts, L.A. ^c	12/70-6/73	19.22	.0315	.0083	.992	Federal Aviation Agency (1974)
Air hijacking deterrences, U.S. ^d	11/67-12/69	69.89	.0390	.0074	.985	New York Times (1967-1973)
Air hijacking deterrences, U.S. ^d	1/70-8/71	83.47	.0282	.0113	.993	New York Times (1967-1973)
Air hijacking deterrences, U.S. ^d	9/71-7/73	151.80	.0576	.0123	.994	New York Times (1967-1973)
Air hijacking deterrences, U.S. ^d	10/68-10/70	35.46	.0274	.0079	.981	New York Times (1967-1973)
Air hijacking deterrences, L.A. ^d	12/70-6/73	18.62	.0178	.0080	.986	New York Times (1967-1973)
Riots, L.A. ^e	1955-1963	188.30	.0031	.0009	.996	Banks (1971)
Guerrilla warfare, L.A. ^e	1955-1963	112.80	.0037	.0009	.996	Banks (1971)
Revolutions, L.A. ^e	1957-1965	78.30	.0025	.0011	.996	Banks (1971)
Purges, L.A. ^e	1958-1966	103.90	.0022	.0013	.999	Banks (1971)
Coups d'etat, L.A. ^e	1958-1966	21.06	.0036	.0010	.993	Banks (1971)
Coups d'etat, Africa ^f	1960-1975	49.01	.0051	.0008	.995	Huff and Lutz (1974), Current History (1972-1975)
Arson, England, 1830 ^g	10/4/30-2/3/31	281.31	.7191	.0513	.997	Hobsbawm and Rudé (1968)
Wage meetings, England, 1830 ^g	10/23/30-12/21/30	168.20	1.2989	.1009	.994	Hobsbawm and Rudé (1968)
Machinery breaking, England, 1830 ^g	11/9/30-12/15/30	249.50	7.4561	.2601	.996	Hobsbawm and Rudé (1968)
Riots, England, 1830 ^g	11/15/30-12/10/30	85.71	2.7540	.2930	.996	Hobsbawm and Rudé (1968)
Civil disorders, U.S., 1966 ^h	3/15/66-10/19/66	23.67	.8627	.0266	.986	U.S. Senate Committee on Gov't. Operations (1968)
Civil disorders, U.S., 1967 ^h	4/1/67-11/21/67	80.12	173.7578	.0724	.943	U.S. Senate Committee on Gov't. Operations (1968)

* This is the upper limit of V, when $q > 0$ $\lim_{t \rightarrow \infty} V = V_0 e^{(c/q)}$

^a This data set is composed of yearly frequencies. Lynchings of blacks occurred before 1882 but that is when the Chicago *Tribune* began collecting and publishing these data. The collection of lynching data was later taken over by the Tuskegee Institute.

Table 1. Continued

^b These outbreaks of vandalism, and swastika paintings on Jewish homes, stores and institutional buildings were nationwide and lasted respectively seven and nine weeks. According to Caplovitz and Rogers (1961) there was no evidence of an organized plot in either of the outbreaks.
^c A hijacking attempt was counted every time a person or a group of persons tried to commandeer an aircraft by threat of harm. The data indicated there were three separate outbreaks of hijacking attempts in the United States and two in Latin America during the August 1967–July 1973 interval.
^d Deterrence attempts are actions taken by combatants which, if successful, might increase potential hijackers' expectations of failure and punishment. These were coded by Miller from the stories on hijacking in the <i>New York Times</i> (1967–1973).
^e Hamblin et al. (1973:126–35) analyzed Banks's data and found, in Latin America, two series of epidemics of various kinds of political violence since World War II. Banks's data for the second series are analyzed here. For definitions see Banks (1971, xv) and Rummel (1963:25–6).
^f Huff and Lutz's data on the diffusion of coups d'état (1960–1972) among 30 contiguous black-ruled countries in Africa were updated through 1975.
^g These data trace an agricultural labor uprising that occurred throughout the east and south of England in the final months of 1830. The data were collected by Hobsbawm and Rudé from all available newspapers as well as public and private records, and are given by date, type of disturbance, place and target.
^h These data are from questionnaires completed by the mayors or city managers of all cities reported to have experienced some sort of riot or civil disturbance. Where information could not be obtained from municipal officials, the data were compiled from local news reports.

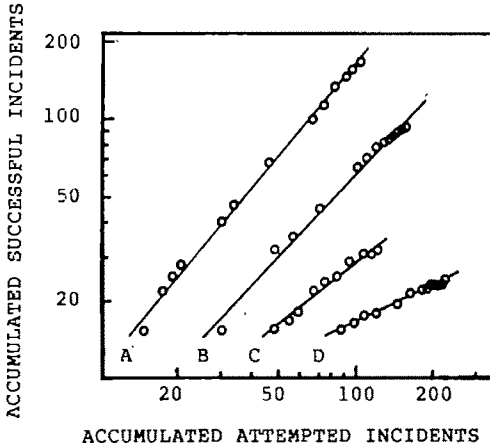
tion and inhibition assumptions analogous to those used in the derivation of equation (8), but with different specifications of the functional form of the inhibition process.

Equation (10) consistently gave the worst fit and for this reason was eliminated from further consideration. Equation (11) gave a poorer fit but was rejected primarily because of the obvious *asymmetry* of some of the data sets which are known to be complete (especially the two anti-Semitic outbreaks, the third U.S. hijacking outbreak, the outbreaks of purges and revolutions in Latin America and the outbreaks of machine breaking and rioting in England). This asymmetry conflicts with the necessary symmetry of the logistic process specified in equation (11) (see fn. 1).

Validity of Assumptions

Continuity of reinforcement contingencies. Equation (8) assumes relatively stable trends in changes in the reinforcement contingencies. One way to display reinforcement conditions in different epidemics is via collective learning curves, by plotting accumulated successes against accumulated attempts. The usual form for such learning curves—individual, organization or collective—is a power function with an exponent somewhat different from 1.0 (cf. Hamblin et al., 1973). The requisite success data are presently available only for the hijacking epidemics and for the coups d'état outbreak in Africa. The appropriate plots are given in Figure 1. Data relationships described by power functions become linear on logarithmic coordinates and that is the case here. Note that the r^2 values are all above .98, and the exponents are quite different from 1.0 except for the African coups. These analyses support the assumptions that reinforcement contingencies are typically not constant and that changes are characterized by continuity.

Premature terminations. Since the model assumes that outbreaks of collective violence are characterized by continuities, discontinuities in reinforcement could prematurely truncate an epidemic and set the conditions for a new one. Examples of early terminations are the



Plot	Outbreak	Exponent	r ²
A.	1st U.S. Hijacking	1.22	.99
B.	Coups d'etat, Africa	1.04	.99
C.	2nd U.S. Hijacking	.79	.98
D.	3rd U.S. Hijacking	.49	.99

Note: On both the ordinate and the abscissa, the plots were positioned for display by multiplying each with a different constant. This does not change the slope or the fit.

Figure 1. Collective Learning Curves for the Three U.S. Air Hijacking Outbreaks (11/67-12/69, 1/70-8/71, 9/71-9/73) and the Black African Coups d'Etat Outbreak (1960-1975)

first and second U.S. epidemics of air hijacking attempts and the first Latin American air hijacking epidemic, all five counter epidemics of deterrence attempts and the Latin American epidemics of riots, guerrilla warfare, and coups d'etat.

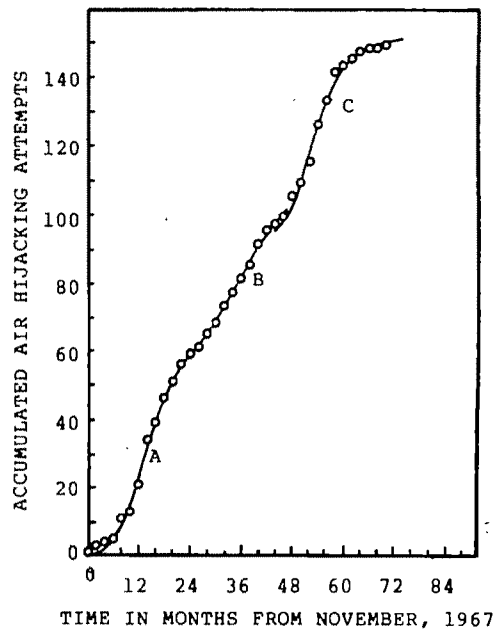
The authors analyzed abstracts of all the stories in the *New York Times* on each hijacking so we have rather detailed knowledge of these events and can, at least, suggest a likely source of the premature truncations in these data.

The first U.S. outbreak mostly involved political migration, hijacking attempts to Cuba which at that time was romanticized as a socialist haven, with imagery of leftist hijackers receiving a hero's welcome. However, that epidemic ended prematurely shortly after six U.S. hijackers returned voluntarily from Cuba to certain prison terms in the United States. They complained bitterly about housing and food. Blacks charged racial discrimina-

tion, and escaped criminals reported they were put into terrible prisons. These stories, given wide coverage in the media, evidently created a discontinuity, enough of a jump in the inhibition process to cut off the outbreak.

The second U.S. outbreak came to a premature end when D. B. Cooper accomplished what appeared to be a successful extortion hijacking by parachuting into the night over Oregon or Washington with \$200,000 from Northwest Airlines. Consequently, the third outbreak began vigorously, with most hijackers attempting extortions and with increased levels of counterviolence.

However, the data suggest overlaps rather than abrupt transitions from one outbreak to another. This is depicted in Figure 2 where the U.S. data are accumu-



Plot	Outbreak	$V_0 e^{c/q}$	c	q
A.	11/67-1/70	70.62	.0677	.0100
B.	1/70-9/71	61.50	.0197	.0066
C.	7/71-9/73	55.91	.0625	.0139

The lines represent the least squares fit of the Gompertz equation to each of the outbreaks with the transition points from one epidemic to another included in both outbreaks.

Figure 2. Air Hijacking Attempts in the U.S. Accumulated over the Entire Period of the Outbreak

lated over the entire period and the transition points are included in both epidemics. The parameters are slightly different from those in Table 1 which were calculated without assumed transitions, and the fit is slightly better.

A discontinuity may also account for the unsatisfactory fit of the model to the 1967 U.S. civil disorder data. In July 1967 a dramatic, massive disorder occurred in Newark, New Jersey. Partly because of the sheer severity of the disorder and partly because it occurred across the river from Manhattan, the U.S. media center, this event received intensive coverage for several days (Spilerman, 1976). This greater than usual media coverage evidently produced a discontinuity, a jump in the instigation process. Over one-half of the 83 disorders in 1967 occurred in the two-week period immediately following the Newark coverage.

The Parameter Estimates

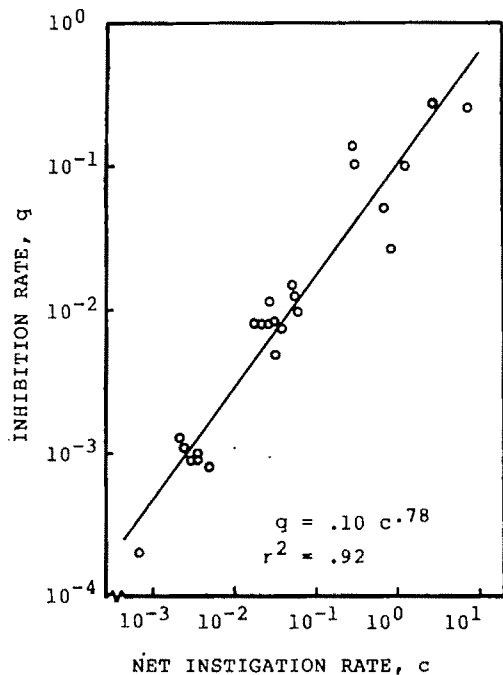
When time is measured using the same unit, as in Table 1 where t is in days, the c and q parameters can be compared across outbreaks. Such comparisons are useful in evaluating the model, since parameters may or may not behave as predicted from the underlying theory. Because of the unsatisfactory fit of the model, the parameters for the 1967 U.S. civil disorder outbreak were not included in the following analyses.

First we consider the expected relationship across violence outbreaks between c , the net instigation rate, and q , the inhibition rate. As noted, past research on the imitation of aggression suggests that instigation rates increase as the benefits of reinforcement to aggressing models increase and inhibition rates increase as their punishment and other costs increase. Violence usually begets violence in equal magnitude—"an eye for an eye and a tooth for a tooth." And, in a violent conflict what is beneficial to one side is usually costly to the other. While the conflict through time continues, the benefits and costs are ordinarily more or less balanced. All of this implies a matching function which in turn implies a positive relationship between the net instigation and inhi-

bition parameters across outbreaks. In general, power functions describe behavior-reinforcement relationships and that is the prediction here.

The c and q parameters from Table 1 are plotted in Figure 3 on logarithmic coordinates. The relationship is linear, indicating the data are described by a power function, and positive, with a least-squares exponent of 0.78. The .78 exponent indicates that outbreaks of violence are generally characterized by undermatching—i.e., relative increases in the instigation rate are greater than the corresponding relative increases in the inhibition rate. The relationship may seem somewhat less than perfect, but an r^2 of .92 is quite high for cross-modality comparisons over such broad continuums of time, culture and violence. In general, these results are very supportive of the model and auxiliary theory.

Second, we consider the relationship between the magnitude of the instigation and inhibition processes and the duration of a violence outbreak. It might be argued



*The solid line represents the least-squares power function given in the figure.

Figure 3. Inhibition Rate (q) for 24 Outbreaks of Collective Violence Plotted on Logarithmic Coordinates by the Net Instigation Rate (c).

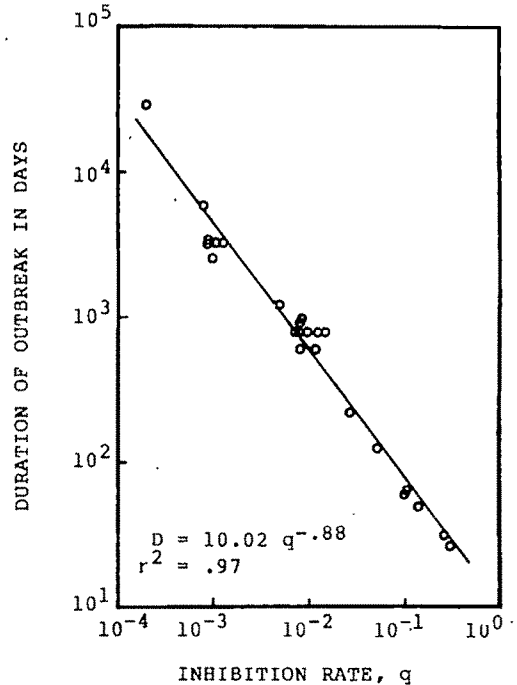
that the lower the net instigation rate and the lower the inhibition rate, the longer an outbreak will last. On the other hand, according to the model incidence of aggression is slowed only by the inhibition rate which functions to terminate the outbreak of violence and thus determine its duration. The hypothesis is that q is the best predictor of duration although c is expected to be highly related because it is a determinant of q . The relationship should be negative and should be described by a power function.

A number of analyses were done, but the one pictured in Figure 4 gives the essential finding: a strong negative relationship between the inhibition rate and the duration of an outbreak. It is described by a power function with an r^2 value of .97 and an exponent of .88. The first order correlation between c , the net instigation rate, and duration was negative and the r^2 value was .90. When c was added as a second independent variable in a multivariate power function, the variance already explained by q was not improved. Other plausible functional forms of these relationships were also tried but their explained variance was substantially lower. Thus, the hypothesis derived from the model is again supported by the data.

Generality

Tilly (1975:514, 519) suggests that the study of collective violence ought to be limited to damage to persons or property by groups of fifty or more in conflict with other such groups, that violence by smaller groups or individuals is either not important or not measurable at a suitable level of accuracy. Contrary to Tilly's (1975) measurement assumption, the size of unit makes no difference in the fit of this model. Since unreliability attenuates relations, this indicates the reliabilities are consistently high for outbreaks with units of all sizes—individuals (e.g., most hijackings), small groups (most hijacking deterrence), crowds (riots) and organizations (revolutionary battles, coups).

The analysis here also questions Tilly's (1975) assumption that small-scale violence is not important enough to study. The data relationships suggest that lesser



*The solid line represents the least-squares power function given in the figure.

Figure 4. Duration in Days of 24 Outbreaks of Collective Violence Plotted on Logarithmic Coordinates by the Inhibition Rate (q)*.

forms of violence often prelude more serious forms as part of an escalation process. For example, the four agrarian labor outbreaks were nested one within another, with the outbreak of arson starting first and ending last, the wage meetings outbreak starting second and ending next to last, then the machine breaking and the riots. The seriousness of the outbreaks may be gauged by the inhibition rates the counterreactions generated: 0.0513 for the arson, 0.1009 for the wage meetings, 0.2610 for the machine breaking and 0.2920 for the riots. The unemployed farm laborers, realizing their earlier strategies were not working, evidently escalated the level of violent conflict in the hope of winning. Hamblin et al. (1973: 129-33) also have published analyses which evidence an escalation process relating the different outbreaks of political violence in Latin America. Snyder (1975:275) has called for the investigation of the life history of collective violence; tracing such escalation processes seems to be a fruitful way to go about it.

Inhibition Effects

The results here suggest that social condemnation can have a very strong inhibitory effect. The perpetrators of the anti-Semitic violence were never caught and, therefore, never subjected to vigilante violence nor prosecuted. Yet those inhibition rates are among the highest exhibited in the outbreaks investigated here. Part of the reason for these high rates seems to have been the rather unanimous and severe condemnation of these acts in the mass media. Memories of Nazi atrocities were still fresh and the outrage against these Nazi imitators was nearly universal. The causes of inhibition are evidently complex and social condemnation may be as important as other counter-reactions.

Certainly counterviolence is often not the most productive way of managing violent outbreaks. When the authorities in Latin America and the United States opted for shootouts with hijackers, about as many passengers and flight personnel were killed as hijackers. In contrast, 10, 16 and 26 percent of the hijackers were talked into surrendering in the first, second and third U.S. outbreaks, respectively. Generally crew members or passengers discussed the options with the hijackers and persuaded them that surrendering was the least noxious alternative.

As noted, most of the violence investigated here involved basic conflicts where the members of both sides were doing physical damage to one another's persons and properties in efforts to settle the issue in their favor. Much of the violence and counterviolence might have been avoided if the conflicts were somehow turned into either legal contests where the facts were considered and the issues adjudicated to effect justice, or political contests where the issues were settled via discussion, debate, compromise and a vote.

Some of the more bizarre outbreaks like the hijackings do not tend to lend themselves to these kinds of solutions, but in most instances that kind of violence is felonious, clearly against the law. In the first U.S. hijacking outbreak, the airlines asked law enforcement agencies to do nothing—to allow the hijacked airliners to

proceed to Cuba without resistance. In the later outbreaks, when the hijackers were forcing pilots to fly them transoceanic and were extorting huge sums, the airliners reversed that policy and law enforcement personnel became quite innovative. As can be seen from the learning curves in Figure 1, they also became much more effective. The hijackings were turned off completely in 1973 when the F.A.A. finally instituted a nonviolent solution: the electronic screening of all passengers for weapons.

CONCLUSIONS

Cumulative distributions of outbreaks of violence are generally sigmoid, some skewed to the right. Our purpose was to develop a model to predict, if possible, the mathematical form of these distributions and to specify the generative processes which could explain these outbreaks. Three models were developed (one detailed), assuming that the benefits and costs to units engaged in a violent outbreak up to any point in time respectively instigated and inhibited subsequent participation by others. It was assumed that a differential equation supported in prior research on logistic models of cultural diffusion is, in fact, a general imitation equation and it was used to predict the form of the instigation process through time. We assumed further that the inhibition process was cumulative and the three models involved different inhibition terms. The model from which the Gompertz function was derived used the imitation equation to describe the inhibition process.

The Gompertz turned out to predict the empirical distributions of violence better than the other equations. The fit was virtually perfect: the median r^2 being .995. There was one deviant case and in that outbreak an essential assumption (i.e., a relatively stable rate of mass media reporting of the violent events) appears to have been badly violated.

The preliminary tests eliminated the alternatives, including application of the logistic diffusion model suggested by Huff and Lutz (1974) and the logistic innovation model suggested by Hamblin et al. (1973). Still, it must be pointed out that the Gom-

pertz might be derived from alternative premises and assumptions. The premises chosen here, however, were empirical generalizations from previous research which were thought to apply in these situations. Also, a number of analyses were done to further evaluate the premises and auxiliary assumptions. Learning curves support the model's assumption that reinforcement contingencies generally exhibited continuity. Early terminations of some of the outbreaks were apparently a function of massive, well published changes in reinforcement contingencies—as would be predicted by auxiliary theory. The inhibition rates were highly related ($r^2 = .92$) to the net instigation rates, a predicted result. The duration of the outbreaks were predicted quite accurately by the value of q ($r^2 = .97$), another predicted result. Because unpredictable variation in parameters can be a prime reason for rejecting a model (cf. Hamblin et al., 1977), these positive results are important corroboration.

This diffusion model has a number of interesting features. First, it appears to be very general: it describes and explains equally well a wide range of violence perpetrated in a number of cultural and historical contexts by units of varying size and type (individuals, small groups, crowds and large organizations). Second, it predicts and explains institutional and dissident violence equally well and thus escapes Firestone's (1974) criticisms that theory and research on violence (a) have focused on dissident violence neglecting institutional violence and (b) have failed to conceptualize a set of processes that account for both.

Finally, the purpose here has been to specify and investigate the processes involved in the timing and contagion of violence, unlike most earlier research which has investigated variables related to the location of violence, the motivation of units engaging in violence, and the social, political and economic conflicts which generate violence. The relative successes here in terms of explained variance, replication and generality suggest these other questions might be profitably investigated in the context of this diffusion model. The constitutive imitation processes as well as

the interactive nature of conflict appear to be too powerful to ignore.

REFERENCES

- Archer, Dane and Rosemary Gartner
1976 "Violent acts and violent times: a comparative approach to postwar homicide rates." *American Sociological Review* 41:937-62.
- Bandura, Albert
1973 *Aggression, A Social Learning Analysis*. Englewood Cliffs: Prentice-Hall.
1977 *Social Learning Theory*. Englewood Cliffs: Prentice-Hall.
- Bandura, Albert, Dorothea Ross and Sheila A. Ross
1963 "Imitation of film-mediated aggressive models." *Journal of Abnormal and Social Psychology* 66:3-11.
- Banks, Arthur S.
1971 *Cross-Polity Time Series Data*. Cambridge, Ma.: MIT Press.
- Caplan, Nathan and Jeffery M. Paige
1968 "A study of ghetto rioters." *Scientific American* 29:15-21.
- Caplovitz, David, and Candace Rogers
1961 *Swastika, 1960: The Epidemic of Anti-Semitic Vandalism in America*. New York: Anti-Defamation League of B'nai B'rith.
- Coleman, James S., Elihu Katz and Herbert Menzel
1966 *Medical Innovation: A Diffusion Study*. Indianapolis: Bobbs-Merrill.
- Davies, James C.
1969 "The J-curve of rising and declining satisfaction as a cause of some great revolutions and a contained rebellion." Pp. 671-709 in Hugh Davis Graham and Ted Robert Gurr (eds.), *Violence in America: A Report to the National Commission on the Causes and Prevention of Violence*. New York: Signet Books.
- Denton, Frank H. and Warren Phillips
1958 "Some patterns in the history of violence." *Journal of Conflict Resolution* 12:182-95.
- Dodd, Stuart C.
1953 "Testing message diffusion in controlled experiments: charting the distance and time factors in the interactance hypothesis." *American Sociological Review* 18:410-6.
1955 "Diffusion is predictable: testing probability models for laws of interaction." *American Sociological Review* 20:392-401.
- Downes, Bryan T.
1958 "Social and political characteristics of riot cities: a comparative study." *Social Science Quarterly* 49:504-20.
- Federal Aviation Administration
1974 "Domestic and foreign aircraft hijackings." *Bulletin: April 2, 1974*. Washington, D.C.: Office of Air Transportation Security.
- Firestone, Joseph M.
1974 "Continuities in the theory of violence: a review." *Journal of Conflict Resolution* 18:117-42.

- Fletcher, R. and M. J. Powell
1963 "A rapidly convergent descent method for minimization." *The Computer Journal* 6:163-8.
- Gilmore, J. Barnard
1971 "Implications for a theory of contagion from an experiment involving model self-sensure." *The Journal of Social Psychology* 85:195-208.
- Griliches, Zvi
1957 "Hybrid corn: an exploration in the economics of technological change." *Econometrics* 25:501-22.
- Gurr, Ted Robert
1968 "Urban disorder: perspectives from the comparative study of civil strife." Pp. 51-68 in Louis H. Massotti and Don R. Bowen (eds.), *Riots and Rebellion: Civil Violence in the Urban Community*. Beverly Hills: Sage.
1970 *Why Men Rebel*. Princeton: Princeton University Press.
- Hamblin, Robert L., Michael Hout, Jerry L. L. Miller and Brian L. Pitcher
1977 "Arms races: a test of two models." *American Sociological Review* 42:338-54.
- Hamblin, Robert L. and Jerry L. L. Miller
1976 "Reinforcement and the origin, rate and extent of cultural diffusion." *Social Forces* 54:743-59.
- Hamblin, Robert L., R. Brooke Jacobsen and Jerry L. L. Miller
1973 *A Mathematical Theory of Social Change*. New York: Wiley.
- Hanratty, Margaret A., Edgar O'Neal and Jefferson I. Sulzer
1972 "Effect of frustration upon imitation of aggression." *Journal of Personality and Social Psychology* 21:30-4.
- Hernes, Gudmund
1972 "The process of entry into first marriage." *American Sociological Review* 37:173-82.
- Hobsbawm, E. J., and George Rudé
1968 *Captain Swing*. New York: Pantheon.
- Huff, David L. and James M. Lutz
1974 "The contagion of political unrest in independent Black Africa." *Economic Geography* 50:352-67.
- Li, Richard P. Y., and William R. Thompson
1975 "The coup contagion hypothesis." *Journal of Conflict Resolution* 19:63-88.
- Liebertson, Stanley and Arnold Silverman
1965 "The precipitants and underlying conditions of race riots." *American Sociological Review* 30:887-98.
- Lupsha, Peter A.
1969 "On theories of urban violence." *Urban Affairs Quarterly* 4:273-96.
- Midlarsky, Manus
1970 "Mathematical models of instability and a theory of diffusion." *International Studies Quarterly* 14:60-84.
- National Advisory Commission on Civil Disorders
1968 *Report*. New York: Bantam.
- Rummel, Rudolph J.
1963 "Dimensions of conflict behavior within and between nations." *General Systems Yearbook* 8:1-50.
- Sears, David G. and T. M. Tomlinson
1968 "Riot ideology in Los Angeles: a study of Negro attitudes." *Social Science Quarterly* 49:485-503.
- Shorter, Edward and Charles Tilly
1974 *Strikes in France: 1930-1968*. London: Cambridge University Press.
- Snyder, David
1975 "Institutional setting and industrial conflict: comparative analyses of France, Italy and the United States." *American Sociological Review* 40:259-78.
- Snyder, David and William R. Kelly
1976 "Industrial violence in Italy, 1878-1903." *American Journal of Sociology* 82:131-62.
1977 "Conflict intensity, media sensitivity and validity of newspaper data." *American Sociological Review* 42:105-23.
- Solaún, Mauricio and Michael A. Quinn
1973 *Sinners and Heretics: the Politics of Military Intervention in Latin America*. Urbana: University of Illinois Press.
- Spiegel, John P.
1971 "Theories of violence: an integrated approach." *International Journal of Group Tensions* 1:77-90.
- Spilerman, Seymour
1970 "The causes of racial disturbances: a comparison of alternative explanations." *American Sociological Review* 35:627-49.
1971 "The causes of racial disturbances: tests of an explanation." *American Sociological Review* 36:427-42.
1976 "Structural characteristics and severity of racial disorders." *American Sociological Review* 41:771-92.
- Tilly, Charles
1975 "Revolutions and collective violence." Pp. 483-555 in Fred I. Greenstein and Nelson W. Polsby (eds.), *Handbook of Political Violence*, Vol. 3. Reading, Ma.: Addison-Wesley.
- Tilly, Charles, Louise Tilly and Richard Tilly
1975 *The Rebellious Century, 1830-1930*. Cambridge, Ma.: Harvard University Press.
- Turner, Ralph H. and Lewis M. Killian
1972 *Collective Behavior*. Englewood Cliffs: Prentice-Hall.
- U.S. Bureau of the Census
1960 *Historical Statistics of the United States, Colonial Times to 1957*. Washington, D.C.: U.S. Government Printing Office.
- U.S. Senate Committee on Government Operations
1968 *Staff Study of Major Riots and Civil Disorders: 1965 through July 31, 1968*. Washington, D.C.: U.S. Government Printing Office.
- Wheeler, Ladd and Anthony R. Caggiula
1966 "The contagion of aggression." *Journal of Experimental Social Psychology* 2:1-10.

WOMEN'S INTERGENERATIONAL OCCUPATIONAL MOBILITY *

RACHEL A. ROSENFELD

McGill University

American Sociological Review 1978, Vol. 43 (February):36-46

Several recent papers have studied women's intergenerational occupational mobility as movement from father's to daughter's occupation. This paper suggests that mother's occupation also is an important dimension of this mobility process. Analysis of data from a national sample of women 30 to 44 in 1967 shows that, when controlling for age and race, a necessary and sufficient model of women's intergenerational occupational mobility includes mother's as well as father's occupation. Whether or not the mother had an occupation outside the home and what occupation she held, given that she was employed, both affect daughter's occupational destination.

Research on women's occupational attainment and mobility has been increasing rapidly during the 1970s. During this decade, several studies have been made of women's intergenerational occupational mobility. DeJong et al. (1971), following Blau and Duncan's (1967) method of analysis, compared the intergenerational occupational mobility patterns of men from Blau and Duncan's sample with the patterns for ever-employed women included in six NORC surveys. On the basis of this comparison, they concluded that mobility patterns did not differ by sex. Tyree and Treas (1974) reanalyzed De Jong et al.'s data, including some of the changes suggested by Havens and Tully (1972) and Rogoff Ramsøy (1973). (See also replies by DeJong et al., 1972; 1973.) Controlling for differences in occupational distribution by sex, Tyree and Treas (1974) did find some interesting differences in the intergenerational occupational mobility patterns of men and women. For instance,

they found that daughters as compared with sons of professional and farmer fathers were more likely to be in white-collar jobs. Tyree and Treas also looked at another type of mobility—from father's to husband's occupation for married women not in the labor force—and found women's intergenerational mobility through marriage more similar to men's than women's intergenerational mobility to their own occupations. Hauser et al. (1974), using samples of men and women more comparable than those used by DeJong et al. (1971), also reported small but significant differences in occupational mobility patterns of women and men. Consistent with Tyree and Treas (1974), Hauser et al. (1974) found sex differences in occupational inheritance (with men more likely than women to enter occupational categories similar to their father's) and accentuation of sex differences in occupational distribution among children of farmers. Supplementing these studies based on data from the 1960s, Featherman and Hauser (1974) examined trends in intergenerational occupational mobility by race and sex in the period 1962-1972. They concluded that while differences between the races and between the sexes in the intergenerational mobility process decreased over the decade, differences by sex were relatively greater than those by race in both 1962 and 1972.

All of these papers define women's intergenerational mobility in the same way as men's is usually defined, i.e., as movement from father's to respondent's

* While doing this research, the author was supported by a University of Wisconsin Knapp Fellowship and by funds granted to the Institute for Research on Poverty at the University of Wisconsin by the Office of Economic Opportunity, pursuant to the provisions of the Economic Opportunity Act of 1964. Computations were carried out in the Center for Demography and Ecology, University of Wisconsin, Madison, which has support from the National Institute for Child Health and Human Development grant No. HD-05876. I wish to thank especially Harry Travis, Robert Hauser, Paul Allison, and Monica Boyd for their helpful comments and suggestions. Naturally, I am responsible for the final content of this paper.

occupational category. Some of the differences by sex found in the above studies could have resulted from the fact that for men, origin is expressed in terms of the occupation of a same-sex parent, while for women it is defined in terms of the occupation of a parent of the opposite sex. The greater similarity to father-son mobility of father-husband than father-daughter mobility is suggestive of this idea. In this paper, I propose that in studying women's intergenerational occupational mobility, mother's occupation be considered as an origin state. Following a discussion of why mother's occupation might be important for an understanding of intergenerational occupational mobility—especially women's—I examine the contribution of mother's as compared with father's occupation to daughter's occupational destination.

Considerations in the Study of Women's Intergenerational Mobility

One can make at least three arguments for including mother's occupation in a study of the structure of intergenerational occupational mobility. First, the inclusion of this variable provides a better measure of family socioeconomic status than father's occupation alone. Second, especially for women, mother's occupation might represent an adult work-role model which affects the occupational choice of her children. Finally, when examining mobility in an occupational structure differentiated by sex, it is necessary to examine intergenerational mobility while holding sex of occupational incumbents constant.

In studying intergenerational occupational mobility, one is usually concerned with the relationship between social origins and current social position. Since occupation is the major indicator of social status in industrial society, social origin is often represented by the occupation of the head of family of origin (usually the father's occupation). Similarly, current social position is represented by respondent's occupation. However, if one thinks of social position only in terms of occupations, one can argue that when the mother works outside the home, father's occupa-

tion alone is not a sufficient indicator of social position of the family. The fact that the mother is earning a wage or salary, regardless of her occupation, generally indicates that she is contributing to the family's financial well-being and to a higher standard of living (U.S. Department of Labor, 1968; Oppenheimer, 1977). In addition, the mother and father may differ somewhat in social position as represented by occupation. It is not unusual, for example, for an employed wife to be in a lower white-collar occupation and for the husband to be in a blue-collar occupation (Watson and Barth, 1964). Although the status distance between employed wives and their husbands is usually not great (Oppenheimer, 1977), still, occupations of both the mother and the father might be used as indicators of the family's general social standing and life style and of the occupation-relevant benefits provided by it to the next generation. Work by Sampson and Rossi (1975) provides evidence that a wife's occupation has some net effect on the family's social status, although less than that of the husband.

When studies of the structure of men's intergenerational occupational mobility use father's occupational category as the origin state, the origin represents both abstract social position and also specific occupational information, role models, and advantages (e.g., a place in a craft union) given to the son. Mothers also may provide important models of and perhaps access to adult work roles, especially for daughters. Research has shown that when the mother has been employed, the daughter is more likely to be employed at any particular time, to be employed more continuously, and to choose and follow less typically female occupations (Almquist and Angrist, 1970; Astin et al., 1971; Baruch, 1974; Hoffman, 1974; Holmstrom, 1973: 199–202; Rapoport and Rapoport, 1971; Siegel and Curtis, 1963; Tangri, 1974). Although such studies do not focus on the actual occupational levels of the mother-daughter generations, they are suggestive of the importance of the mother's occupation for that of the daughter and point to the need to study this directly.

The nature of the occupational structure suggests a third reason for including mother's occupation in a study of female intergenerational mobility. The occupational structure is highly differentiated by sex. Some occupations are more likely than others to attract and/or be filled by women. Women are more likely than men to be concentrated in occupations such as clerical and semiprofessional occupations which require relatively high levels of education, receive relatively high levels of prestige or status (since they are white-collar occupations), but in which wages are relatively low (Oppenheimer, 1970; Treiman and Terrell, 1975b). The greater probabilities for daughter than for sons of entering low level white-collar occupations, which exist even after controlling for sex differences in occupational destinations (Tyree and Treas, 1974), probably reflect the differences in occupational opportunities open to and sought by women and men.¹ Since the effects of sex are held constant, examining mobility from mother's to daughter's occupation might be helpful in answering questions about the relationship between parent's and child's occupational locations.

Sociologists generally study intergenerational occupational mobility because of a concern with the openness of the occupational structure, with the opportunities available to people from different backgrounds. Comparisons of mobility patterns by sex, then, are comparisons of the opportunities open to men and women. Including mother's occupation in the study of intergenerational mobility, especially that of women, might both characterize social origin more fully and also point to some of the mechanisms by which status is transmitted across generations.

Given the above arguments for includ-

ing mother's occupation as an aspect of the origin state when examining women's intergenerational occupational mobility, I now turn to an empirical test of the assertion that both mother's and father's occupations are important dimensions of women's intergenerational mobility. The data and categorization used in constructing a five-dimensional mobility table—father's occupation by mother's occupation, by daughter's occupation, by daughter's race, by daughter's age—are described. Then log-linear models are used to analyze this table.

The analysis will first determine which dimensions are really needed to describe adequately women's intergenerational occupational mobility. In particular, are there sufficient models of women's intergenerational mobility which do not include mother's occupation? It is possible for a log-linear model adequately to predict cell frequencies, i.e., be sufficient, while excluding variables which would contribute significantly to explaining the associations in the table under consideration. The purpose of the second part of the analysis, then, is to test whether mother's occupation has a significant effect, even given sufficient models which do not include this dimension. The effect of mother's occupation will be examined to determine whether it is actually level of mother's occupation or simply mother's employment which affects the daughter's occupational destination. Finally, the relative importance of the effect of mother's as compared with father's occupation on daughter's occupation will be described.²

Design

Steinmetz (1974: 115–6) notes that the National Longitudinal Surveys of Work

¹ Of course, since there are differences by sex in placement within occupational categories, similarity in patterns of intergenerational mobility across broad occupational categories would not necessarily mean that women have the same opportunities as men. One reviewer suggested that in comparisons of mobility patterns by sex, some combination of occupational category and sex-type be used to measure location within the occupational structure. For a study of the effects of sex-type of occupation on intragenerational status mobility, see Wolf and Rosenfeld (1978).

² The next logical step would be to investigate the effects of occupations of both parents on men's occupational movement so that those results could be compared with the results obtained for women. Data were not immediately available to test these effects. The differences discovered in mobility patterns by sex in the research using father's occupation alone as an indicator of occupational origin, the literature on role models provided by same-sex parent, and the observed differences in the occupational structure by sex at a given time all suggest that the effect of mother's occupation on son's attainment is probably less than on daughter's.

Experience, carried out for the Manpower Administration of the U.S. Department of Labor, are unusual in collecting occupational data on respondents' mothers as well as fathers. This paper uses data from the survey of mature women, mature women being one of the four subsets of the population included in the Surveys.³

To obtain initial information about the labor force experience of mature women, the Bureau of the Census conducted personal interviews in May, 1967, with a national probability sample of the civilian, noninstitutionalized population of women who were 30 to 44 years of age in April, 1967. The total sample consisted of 5,083 women; 3,606 of whom were white. (See Shea et al., 1970, for further description of the sample and the sampling process.⁴)

Within this sample, I will be focusing on those women who came from families in which there were two parents (or stepparents) to provide adult work-role models when the respondent was 15, and who were themselves in a civilian occupation at some time before they were interviewed in 1967. There were 3,472 such respondents.

Occupational origin is operationalized as the occupations or nonemployment of both parents when the respondent was 15 years of age. Occupational destination is the woman's current or last occupation.⁵

³ The other three cohorts were men 45 to 59 years of age, and young men and women 14 to 24 years of age. The later surveys of the younger cohorts, not available when this paper was in progress, might be used to compare intergenerational mobility patterns by sex, as suggested in fn. 2.

⁴ As Treiman and Terrell (1975a: 175) also noted, the number of whites and of nonwhites reported in Shea et al. (1970) is incorrect.

⁵ There is a continuing debate over how to characterize women's social position. Underlying an interest in women's mobility through marriage is often the belief that "in the United States and indeed much of the industrialized world the main determinant of a man's social status is his occupation, and the main determinant for a woman is the social position of her husband" (Chase, 1975: 484). In this paper, a woman's social position is characterized by her own current or last place within the occupational structure.

There are at least two other ways in which destination state might have been considered: (1) as the current occupation of employed women and (2) as including housewife as an occupational category (as is done for mothers). The first alternative has been

These origin and destination occupations are categorized as follows:

- (1) professional, technical, managerial;
- (2) clerical, sales;
- (3) crafts, operative, nonfarm labor;
- (4) service, including private household service;
- (5) farm, including farm labor;
- (6) not employed (origin only).

This categorization limits the number of categories so that the mobility table will not contain an excessive number of empty cells. At the same time, it maintains crucial distinctions among occupational groups. One can assume that most of the mothers for whom no occupation was reported were housewives when their daughters were 15.

A word of explanation is in order here as to why occupations are assigned to broad occupational categories rather than being assigned some measure of prestige or socioeconomic status. Occupational sex segregation does not follow socioeconomic lines closely. In contrast with findings from analysis using occupational categories, comparisons of the socioeconomic status or prestige attainment of women and men find little difference by sex either in mean attainment or in the effects of social origins on attainment (see Featherman and Hauser, 1976; McClenendon, 1976; and Treiman and Terrell, 1975a). Use of prestige or status scores seems to hide sex differences in occupational location and mobility. Since the purpose here is to consider the place of mother's as compared with father's occu-

used in studies of occupational destination (e.g., Treiman and Terrell, 1975a). However, this definition of women's occupational position neglects the labor force experience of many women. Most of the women surveyed had been employed at some time since leaving school, although many were not currently employed. By considering last occupation held, even if this is not the current occupation, one gets a better idea of where women go when they are in the labor force. (A similar argument could be applied to mother's occupation as well. However, information is available only on the occupation held by the mother when the respondent was 15.) Including "housewife" as a category was another possibility. It is in the labor force, however, that occupational mobility is usually thought to occur. (But see Featherman and Hauser, 1974.)

pation in the structure of intergenerational occupational mobility, occupational categories are used rather than measures of the location of occupations in a hierarchy of prestige or status. Ideally, however, finer categories would have been used.

There is a further issue involved here: describing the full range of work done by women. Although there have been attempts to measure the prestige of housewife (e.g., Bose, 1973), such a prestige score confuses prestige rewards obtained within the occupational structure with prestige obtained otherwise. The problems caused by the lack of a prestige or status score for the occupation of housewife are illustrated by Treiman and Terrell's (1975a) study. Using data from the same general sample used here, they did include mother's occupation in their model of employed women's attainment. Since they were focusing on prestige attainment, rather than on occupational mobility per se, they coded detailed occupations according to a scale of occupational prestige. Lacking a code for housewife, an occupation held by a large proportion of the mothers, data on mother's occupation were considered missing for about 70% of the cases in their subsample. When mother's occupation was not reported because the mother had been a housewife (or in some cases because the mother was unemployed or no mother was present in the respondent's household at the time the respondent was 15), they coded this variable as the mean for reported cases. Even though this coding decreased the possibility of finding a signifi-

cant effect by usual tests, they found a statistically significant effect of mother's occupation on the daughter's occupation. In this paper, occupations are grouped into categories which include that of "no employment" for mothers.

Occupational distributions do vary by race. Therefore, one might expect that race would have some effect on the occupational mobility process. Birth cohort, which represents the common experience of a group of individuals, is another factor besides occupational origins which might affect the occupational mobility experienced by women. Ages in the survey of mature women range from 30 to 44 in 1967. Women born in 1937 and leaving school in the 1950s (a time when women were told that their role was to remain in the home) might have mobility patterns different from women born in 1923 and leaving school during World War II (when women were called into the labor force). In addition, age itself might reflect factors such as length of time in the labor force and opportunity to achieve higher occupations. Therefore, race and age were included in the mobility table as control dimensions. Race is coded dichotomously as white/nonwhite. Age is divided into three five-year groups: 30-34, 35-39, 40-44.

Each respondent's data were weighted such that the adjusted data represent the national distribution for women 30-44 meeting the criteria for inclusion in this sample (see Shea et al., 1970: App. B). The percentage distributions for categories of parental occupations, respondent's occupation, race and age are shown in Table 1.

Table 1. Marginal Percentages for Parental Occupations, Respondent's Occupation, Race and Age^a

Variable	Categories ^b						Total
	1	2	3	4	5	6	
Father's occupation	19.2	6.6	46.0	4.2	23.0	1.0	100.0
Mother's occupation	3.5	5.5	8.1	8.5	3.4	71.0	100.0
Respondent's occupation	16.5	46.7	18.0	16.2	2.6		100.0
Race	91.2	8.8					100.0
Age	31.6	33.5	34.9				100.0

N = 3472

^a Percentages are based on weighted frequencies.

^b Occupational categories are: 1=professional, technical, managerial; 2=clerical, sales; 3=crafts, operative, labor; 4=service, including private household; 5=farm; 6=not employed. Race categories are: 1=white; 2=nonwhite. Age categories are: 1=30-34 years of age in 1967; 2=35-39 in 1967; 3=40-44 in 1967.

Table 2. Models of Intergenerational Occupational Mobility of Women

Model Number	Marginals ^a Fitted	χ^2_{LR}	Adjusted df	p- value	Index of dissimilarity between expected and observed frequencies
1	{FMAR ^b } {D}	1215.21	580	<.001	19.42
2	{DF}	745.12	560	<.001	13.17
3	{DM}	903.33	560	<.001	16.55
4	{DR}	988.87	576	<.001	16.56
5	{DA}	1192.84	572	<.001	19.30
6	{DF} {DM}	577.02	540	>.10	11.24
7	{DF} {DR}	570.24	556	>.10	10.72
8	{DF} {DA}	724.05	552	<.001	12.88
9	{DM} {DR}	743.36	556	<.001	14.33
10	{DM} {DA}	884.10	552	<.001	16.38
11	{DR} {DA}	966.04	568	<.001	16.63
12	{DF} {DM} {DR}	439.25	536	>.50	9.49
13	{DF} {DA} {DR}	548.94	548	>.10	10.00
14	{DM} {DR} {DA}	723.94	548	<.001	14.24
15	{DF} {DM} {DA}	557.88	532	>.10	10.65
16	{DF} {DM} {DA} {DR}	420.01	528	>.50	8.61

^a F = Father's occupation; M = Mother's occupation; D = Respondent's occupation; R = Race; A = Age.

^b Included in all other models, also.

Models of Women's Intergenerational Occupational Mobility

Log-linear models are used here to analyze women's intergenerational occupational mobility. These models are particularly suited to the analysis of dichotomous or dichotomized and polytomous or polytomized variables. (See Goodman, 1970; 1972a; 1972b; 1973; Fienberg, 1970; and Bishop et al., 1975, for detailed descriptions of these models and their interpretation.)

The first task in the analysis is to find sufficient models of women's intergenerational occupational mobility. Essentially, the aim here is to discover how little of the information in the five-dimensional mobility table is needed to describe the associations in it. Of particular importance is whether there are sufficient models which do not take into account the distribution of daughter's by mother's occupation. In Table 2, various models are tested to see how well they account for women's intergenerational occupational mobility.

Before the discussion of specific models, Table 2 requires some general explanation. The letters in brackets at the left of the table indicate which marginals or combinations of marginals have been fitted for that model. Notice that here daughter's occupation is conceptualized as a dependent variable to be explained by

parent's occupations, controlling for age and race. The marginal table of associations among the independent variables is therefore considered fixed and is fitted under every model tested. The chi-square value indicates how well a given model reproduces the frequencies in the classification of daughter's occupation by father's occupation, by mother's occupation, by race, by age. A significant chi-square value indicates that the frequencies expected under that model differ significantly from those actually observed and that other marginals of the tables have to be fitted in order to reproduce adequately the observed cell frequencies.⁶ The index of dissimilarity (Δ) between observed and expected values gives another indication of how well a model fits the data. (See

⁶ The likelihood ratio chi-square is used here:

$$\chi^2_{LR} = 2 \sum_i \sum_j \sum_k f_{ijk} \log_e \left(\frac{f_{ijk}}{F_{ijk}} \right)$$

where f_{ijk} is the observed frequency in cell ijk of (for the sake of example) a three-dimensional contingency table, and F_{ijk} is the frequency in that cell expected under a given model.

The advantages of using a likelihood ratio chi-square rather than the usual Pearson chi-square in the analysis of contingency tables are (1) the likelihood ratio chi-square can be conveniently partitioned and (2) the χ^2_{LR} is the statistic which is minimized by maximum likelihood estimates (Bishop et al., 1975:125-30).

Taeuber and Taeuber, 1965, for a discussion of this index.)

Model 1 represents the situation in which the occupational distribution of daughters is independent of the parents' occupational distributions as well as of the distributions across categories of race and age. Models 2 through 5 assume that the distribution of daughter's occupation depends on only one of the four other variables and that this dependence is the same within various subsets defined by the other categories. Models 6 through 16 include various combinations of these main effects.

Tests of the fit of the models in Table 2 indicate that the respondent's occupation is not independent of the other four dimensions of the mobility table (see model 1) and that the control variables are not sufficient, alone or in combination, to account for the observed patterns in the mobility table (models 4, 5, 11). The models which include effects of both mother's occupation and father's occupation (models 6, 12, 15, 16) provide estimates of cell frequencies which do not differ significantly from those actually observed. One could interpret this as a confirmation that at least both mother's and father's occupations are needed to describe the origin state in analysis of women's occupational mobility across generations. However, the models including effects of both father's occupation and of race, though not mother's occupation (models 7, 13), also fit the data well. The same is not true for the models including effects of both mother's occupation and race but not father's occupation (models 9, 14). It could be that either race or mother's occupation adds enough additional information about family social position to the information contained by father's occupation to determine daughter's occupation.⁷

⁷ The proper significance level to use here is problematic. In Table 2, the degrees of freedom have been adjusted downward because of the presence (as a result of using a sample) of zero marginals, which put additional constraints on the fitting process. The *p*-values given are the conventional ones for these adjusted degrees of freedom and the unadjusted chi-square values. No adjustment has been made for the efficiency of the sampling design (the design is not a simple random sample) or for the use of multiple significance tests (Goodman, 1969:11).

There are models sufficient to reproduce the women's intergenerational occupational mobility table which do not include the association between mother's and daughter's occupation. What is important, however, is not simply whether a model fits the data, but whether some parameters not included would significantly improve the fit of expected to observed frequencies. In particular, does mother's occupation have a significant effect on daughter's occupation even after including father's occupation and race in the model? To test whether additional parameters are significant, one compares the chi-square values for models with and without the parameters to be tested. The difference between the chi-square values is itself distributed as chi-square with degrees of freedom equal to the differences in degree of freedom of the models being compared. If the chi-square for the difference is significant, then the additional parameters do contribute significantly to the prediction of the cell frequencies.

By comparing models from Table 2 which do and do not include mother's occupation, one tests whether mother's occupation has a significant effect, given the initially included variables. As can be seen from Table 3, whether the baseline model is the model of independence or the model including all main effects, the model containing the effects of both father's occupation and mother's occupation, as well as those of the control characteristics, is necessary to explain the relationships in women's intergenerational mobility table.⁸ Despite the fact that models containing only father's occupation and race fit the data rather well, the dimension of mother's occupation adds significantly to prediction of the daughter's occupation.

Considering the importance which the occupational choice and the family literature (e.g., Hoffman, 1974) give to the effect of maternal employment outside the home per se, it is important to determine whether the effect of mother's occupation found here is actually the result of level of

⁸ Tests were made for the possible significance of interactions between father's and mother's occupations and of age or race with parent's occupation. There were no significant interaction effects.

Table 3. Testing for Increments to Models

Models ^a Compared	Parameter(s) ^b Tested	χ^2_{LR}	df	p
a. Testing for increments to independence model: Zero-order χ^2 's				
2 vs. 1	DF	470.09	20	<.001
3 vs. 1	DM	311.88	20	<.001
	(1) occupation of mother (except housewife)	292.59	16	
	(2) mother employed/ housewife	19.29	4	
4 vs. 1	DR	226.34	4	<.001
5 vs. 1	DA	22.37	8	.004
b. Testing for contributions to model containing all main effects				
14 vs. 16	DF	303.93	20	<.001
13 vs. 16	DM	128.93	20	<.001
15 vs. 16	DR	137.87	4	<.001
12 vs. 16	DA	19.24	8	.014

^a Model numbers refer to Table 2.

^b F = Father's occupation; M = Mother's occupation; D = Respondent's occupation; A = Age; R = Race.

mother's occupation or is the effect of the simple contrast between mother's employment outside the home or her nonemployment. Have we simply found that it matters for daughter's occupation whether or not the mother worked outside the home or have we really found an association between mother's and daughter's occupational locations?

This question can be answered by partitioning the zero-order chi-square from the comparison of model 1 with model 3 into that part due to the contrast between mother's employment vs. lack of employment and that part due to the five levels of mothers' reported occupations (see Goodman, 1968; Duncan, 1975). Models 1 and 3 were estimated and compared for two subtables of the original table: one in which mother's occupation was dichotomized as employed vs. not employed, and one in which all cases in which mother's occupation was not reported were eliminated. These comparisons are shown as lines (1) and (2) in panel a. of Table 3. The chi-square attributable to the contrast between employed and not employed mothers is 19.29 with 4 degrees of freedom. That attributable to the actual occupation of the employed mothers is 292.59 with 16 degrees of freedom. Both parts are significant. One can conclude that there is a relationship between mother's occupation per se and daughter's occupation and that the effect of mother's

occupation (for all six categories) is more than an effect of mother's employment or lack of employment.

The tests of models have shown that while there are models which sufficiently account for women's intergenerational occupational mobility which do not include mother's occupation, mother's occupation has a significant effect even given father's occupation, race, and age. The relative importance of mother's and father's occupations within the model including all main effects can be described using a statistic suggested by Goodman (e.g., 1972a: 42-4):

$$\frac{\chi^2(H'') - \chi^2(H')}{\chi^2(H'')} \quad (1)$$

where the parameters in the model H'' are included among the parameters in the model H' . Table 4 presents the relative reduction in chi-square made separately and as a group by the variables father's occupation, mother's occupation, race, and age. When the data include the cases in which the mothers were not employed, father's occupation brings about the greatest relative improvement in fit, given that the other variables were included in the model. The coefficient for mother's occupation is smaller and about equal to that for race (see Table 4: col. A).

The analysis of the effect of mother's occupation showed that the occupational level of employed mothers does signifi-

cantly affect the occupational location of daughters. To get an idea of the relative importance of father's and mother's occupation when both parents were (with a few exceptions among the fathers) employed outside the home, the relative reductions in chi-square were calculated excluding the cases in which mothers had been housewives when their daughters were 15 (see Table 4: col. B). Mother's occupation now brings about the greatest relative reduction in chi-square.

When we are concerned with a group of women, the majority of whose mothers were housewives, the distribution of father's occupation contributes most to explaining the distribution of the daughters across occupational categories. When the mothers held occupations outside the home, at least when the daughters were 15, then the distribution of these mothers over occupational categories contributes more to predicting the occupational distribution of the daughters than does the fathers' distributions. This is support for the idea that the role model and occupational knowledge, and additional financial resources offered by the employed mother are even more important than the family's social position (as represented by the father's occupation alone) for the occupational destination of the daughter. When the mother does not hold a job outside the home, then father's occupation, perhaps an indicator of family life style and occupation-relevant benefits received by the daughter, has a stronger relationship with daughter's occupation.

Nothing has been said yet about differences in the associations found between father's and daughter's and mother's and daughter's occupation. Some idea as to the nature of these differences can be gained by examining estimates of the weights for predicting cell frequencies associated with the categories and combinations of categories included in the model found necessary and sufficient here, model 16 of Table 2. The parameter estimates for DF and DM combinations of categories reflect net association between the two pairs of variables. These estimates (not shown), calculated for all respondents and for only those reporting that their mothers were employed, suggest a somewhat greater inheritance of daughter's occupation from the mother than from the father, especially of clerical/sales and farm occupation, and a greater association between professional and clerical/sales occupations for mothers and daughters as compared with fathers and daughters. These differences are consistent with the arguments developed so far. The greater inheritance of clerical/sales occupations and the greater association between professional and clerical occupations could reflect differences in the occupational locations of men and women within and across occupational categories. The greater inheritance of the farm category might indicate that mother's employment in a farm occupation is a particularly strong indicator of family social position or that in a farm setting, the role model given by the mother is especially

Table 4. Relative Reduction to Chi-Square with Addition of Parameters

Relative reduction in chi-square when:	A ^a	B ^b
{DF} ^c added to model containing all other main effects	303.93/723.94=.42	46.47/350.28=.13
{DM} added to model containing all other main effects	128.93/548.94=.23	118.37/422.18=.28
{DR} added to model containing all other main effects	137.87/557.88=.25	41.36/345.17=.12
{DA} added to model containing all other main effects	19.24/439.25=.04	9.49/313.30=.03
all main effects added to independence model	795.20/1215.21=.65	392.78/696.59=.56

^a Calculated from data including cases in which mother was not employed.

^b Calculated from data excluding cases in which mother was not employed.

^c D = Respondent's occupation; F = Father's Occupation; M = Mother's occupation; R = Race; A = Age.

important for the daughter's occupational future. These differences in the associations of mother's and daughter's and father's and daughter's occupations are also generally complementary to the differences by sex in mobility from father's occupation. This examination of differences in patterns of mobility is only tentative, however, since the standard errors of the estimates are large.

Conclusion

Here a necessary and sufficient model of women's intergenerational occupational mobility was found to include the effects of mother's occupation as well as of father's occupation, age, and race on daughter's occupation. Mother's occupation is a significant dimension of women's intergenerational occupational mobility. Whether or not the mother worked outside the home and what occupation she held, given that she was employed, affect daughter's location. When only those cases are considered in which the mother had been employed at the time the daughter was 15, the effect of mother's occupation is relatively more important than that of the father in predicting daughter's occupational destination. In other words, mother's work matters for daughter's occupational location. This same conclusion has been reached with respect to girl's and young women's occupational aspirations and with respect to the occupational attainment of certain selected subgroups (primarily college graduates). The findings here allow generalization to the occupational mobility process for adult women.

The study here adds to the growing body of research on women's occupational attainment and mobility. Most of the previous studies in this area have focused on women's mobility from their fathers' occupational positions or on the process of socioeconomic attainment. Here emphasis was on the place of mother's occupation in women's intergenerational occupational mobility. The results of this analysis indicate that, especially as more women find employment outside the home, those concerned with the process by which women find their places within the occupational struc-

ture should consider the impact of mother's as well as father's occupation on daughter's occupation. Further, it is possible that mother's occupation is associated with son's occupation as well as daughter's. This possibility bears investigation.

REFERENCES

- Almquist, Elizabeth M. and Shirley Angrist
1970 "Career saliency and atypicality of occupational choice among college women." *Journal of Marriage and the Family* 32:242-9.
- Astin, Helen, Nancy Suniewick and Susan Dweck
1971 *Women: A Bibliography on Their Education and Careers*. Washington, D.C.: Human Service Press.
- Baruch, Grace K.
1974 "Maternal career-orientation as related to parental identification in college women." *Journal of Vocational Behavior* 4:173-80.
- Bishop, Yvonne M., Stephen E. Fienberg and Paul Holland
1975 *Discrete Multivariate Analysis*. Cambridge, Ma.: MIT Press.
- Blau, Peter M. and Otis Dudley Duncan
1967 *The American Occupational Structure*. New York: Wiley.
- Bose, Christine E.
1973 *Jobs and Gender: Sex and Occupational Prestige*. Baltimore: Center for Metropolitan Planning and Research, Johns Hopkins University.
- Chase, Ivan
1975 "A comparison of men's and women's intergenerational mobility in the United States." *American Sociological Review* 40:483-505.
- DeJong, Peter Y., Milton J. Brawer and Stanley Robin
1971 "Patterns of female intergenerational occupational mobility: a comparison with male patterns of intergenerational occupational mobility." *American Sociological Review* 36:1033-42.
- 1972 "Reply to Havens and Tully." *American Sociological Review* 37:777-9.
- 1973 "Patterns of female intergenerational occupational mobility: response to Ramsøy." *American Sociological Review* 38:807-9.
- Duncan, Otis Dudley
1975 "Measuring social change via replication of surveys." Pp. 105-27 in Kenneth Land and Seymour Spilerman (eds.), *Social Indicators Models*. New York: Russell Sage Foundation.
- Featherman, David L. and Robert M. Hauser
1974 "Trends in occupational mobility by race and sex in the United States, 1962-1972." Working Paper 74-25, Center for Demography and Ecology, University of Wisconsin, Madison.
- 1976 "Sexual inequalities and socioeconomic achievement in the U.S., 1962-73." *American Sociological Review* 41:462-83.

- Fienberg, Stephen E.
1970 "The analysis of multidimensional contingency tables." *Ecology* 51:419-33.
- Goodman, Leo
1968 "The analysis of cross-classified data: independence, quasi-independence, and interactions in contingency tables with or without missing entries." *Journal of the American Statistical Association* 63:1091-131.
1969 "How to ransack social mobility tables and other kinds of cross classification tables." *American Journal of Sociology* 75:1-40.
1970 "The multivariate analysis of qualitative data: interactions among multiple classifications." *Journal of the American Statistical Association* 65:226-56.
1972a "A modified multiple regression approach to the analysis of dichotomous variables." *American Sociological Review* 37:28-46.
1972b "A general model for the analysis of surveys." *American Journal of Sociology* 77:1035-86.
1973 "Causal analysis of data from panel studies and other kinds of surveys." *American Journal of Sociology* 78:1135-91.
- Hauser, Robert M., David L. Featherman, and Dennis P. Hogan
1974 "Race and sex in the structure of occupational mobility in the United States, 1962." Working Paper 74-26, Center for Demography and Ecology, University of Wisconsin, Madison.
- Havens, Elizabeth M. and Judy Corder Tully
1972 "Female intergenerational mobility: comparison of patterns?" *American Sociological Review* 37:774-7.
- Hoffman, Louis Wladis
1974 "Effects on the child." Pp. 126-66 in Louis Wladis Hoffman and F. Ivan Nye (eds), *Working Mothers*. San Francisco: Jossey-Bass.
- Holstrom, Lynda Lytle
1973 *The Two-Career Family*. Cambridge, Ma.: Schenkman.
- McClendon, McKee
1976 "The occupational status attainment processes of males and females." *American Sociological Review* 41:52-64.
- Oppenheimer, Valerie
1970 *The Female Labor Force in the United States: Demographic and Economic Factors Governing its Growth and Changing Composition*. Population Monograph Series, No. 5. University of California, Berkeley.
1977 "The sociology of women's economic role in the family." *American Sociological Review* 42:387-406.
- Rapoport, Rhona and Robert N. Rapoport
1971 "Early and later experiences as determinants of adult behavior: married women's family and career patterns." *British Journal of Sociology* 22:16-30.
- Rogoff Ramsøy, Natalie
1973 "Patterns of female intergenerational occupational mobility: a comment." *American Sociological Review* 38:806-7.
- Sampson, William A. and Peter H. Rossi
1975 "Race and family social standing." *American Sociological Review* 40:201-14.
- Shea, John R., Ruth S. Spitz and Frederick A. Zeller
1970 *Dual Careers: A Longitudinal Study of Labor Market Experience of Women*, Vol. 1. Manpower Research Monograph No. 21. Washington, D.C.: U.S. Government Printing Office.
- Siegel, Alberta Engvall and Elizabeth Ann Curtis
1963 "Familial correlates of orientation toward future employment among college women." *Journal of Educational Psychology* 54:33-7.
- Steinmetz, Suzanne
1974 "The sexual context of social research." *The American Sociologist* 9:111-6.
- Taeuber, Karl E. and Alma F. Taeuber
1965 *Negroes in Cities: Residential Segregation and Neighborhood Changes*. Chicago: Aldine.
- Tangri, Sandra Swartz
1974 *Effects of Background, Personality, College and Post-College Experience on Women's Post-Graduate Employment*. Washington, D.C.: U.S. Commission on Civil Rights, U.S. Department of Labor, Manpower Administration.
- Treiman, Donald J. and Kermit Terrell
1975a "Sex and the process of status attainment: a comparison of working women and men." *American Sociological Review* 40:174-200.
1975b "Women, work, and wages—trends in the female occupational structure." Pp. 157-99 in Kenneth C. Land and Seymour Spilerman (eds.), *Social Indicator Models*. New York: Russell Sage Foundation.
- Tyree, Andrea and Judith Treas
1974 "The occupational and marital mobility of women." *American Sociological Review* 39:293-302.
- U.S. Department of Labor, Wage and Labor Standards Administration, Women's Bureau
[1968] "Working wives—their contribution to family income." Pp. 194-7 in Nona Glaser-Malbin and Helen Youngelson Waehrer (eds.), *Woman in a Man Made World*. Chicago: Rand McNally.
- Watson, Walter B. and Ernest A. T. Barth
1964 "Questionable assumptions in the theory of social stratification." *Pacific Sociological Review* 7:10-6.
- Wolf, Wendy C. and Rachel A. Rosenfeld
1978 "Sex structure of occupations and job mobility." *Social Forces*. In press.

CURRICULUM TRACKING AND EDUCATIONAL STRATIFICATION: SOME FURTHER EVIDENCE*

KARL L. ALEXANDER, MARTHA COOK AND EDWARD L. MCDILL

Johns Hopkins University

American Sociological Review 1978, Vol. 43 (February):47-66

Utilizing longitudinal survey data from a subsample of a national sample of youth contacted in the ninth, eleventh, and twelfth grades, an elaborated "school process" model is evaluated to determine the differential antecedents and consequences of high school curriculum placement. The effects of curriculum differentiation on academic achievements (both relative and absolute), educational goals, two behaviors relating to educational goals (application to and acceptance by a college), and social supports for educational attainment are considered. Precurriculum controls at the junior high level on these outcomes provide a stringent assessment of tracking effects not available in prior research. Socioeconomic characteristics of students influence curriculum enrollment in high school almost totally through their effects on achievements, goals, and encouragement during junior high school. Net of numerous pre-enrollment control variables, curriculum placement has important effects on educational outcomes in the junior and senior years; it serves both to mediate the effects of prior variables in the model and to contribute uniquely to the explanation of these outcomes. Curriculum assignments and consequences revealed in the analysis are interpreted in light of functional vs. conflict theories of educational stratification, and it is concluded that neither provides an entirely adequate explanation of such differentiation.

In his 1976 presidential address before the Southern Sociological Society, Alan Kerckhoff (1976) offered a critical appraisal of recent status attainment research. He quite correctly observed that this literature, especially the so-called "school process" modeling, generally has neglected organizational and structural arrangements which may constrain educational opportunities and outcomes. This imbalance reflects the enduring influence of the initial "Wisconsin" social-psychological models of educational stratification (Sewell et al., 1970; Sewell et al., 1969). In attempting to understand how status origins affect educational and socioeconomic attainments, these studies focused on mechanisms of socialization and on their importance in shaping students' motivations and values.

As a complement to the social-psychological perspective of the Wisconsin framework, Kerckhoff advocates con-

sideration of structural constraints in the social organization of schooling which may condition educational outcomes entirely independent of the kinds of interpersonal and subjective processes so important to the Wisconsin model. The few studies to have included selection and allocation mechanisms in such models (e.g., Alexander and Eckland, 1975; Alexander and McDill, 1976; Heyns, 1974; Hauser et al., 1976; Rosenbaum, 1975) buttress Kerckhoff's position. These studies, focusing on curriculum differentiation, examine how tract membership provides access to various educational resources and promotes or retards achievement. Whether one is enrolled in a college or a noncollege track has been found to be of considerable consequence across a broad range of outcomes, including academic performance, encouragement from significant others, educational goals, and self-conceptions of competence.

The study by Alexander and McDill (1976) is the most recent and comprehensive of these inquiries. Since the present paper builds upon their effort in several respects, we first briefly review their findings and then develop the ways in which this report extends and refines their work.

* Paper presented at the annual meetings of the American Sociological Association, Chicago, 1977. The research reported herein was supported by the Education and Work Group of the National Institute of Education. We would like to thank Larry J. Griffin, and two anonymous reviewers for their helpful suggestions.

Alexander and McDill considered the importance of selected social background and demographic characteristics for enrollment in a college-preparatory program and the consequences of such enrollment for a number of important educational outcomes. In terms of curriculum sorting, perhaps their most interesting finding was the substantial influence of status origins, which was almost as important as measured ability in determining track placement. Curriculum membership itself was found to exert appreciable influence on *all* subsequent outcomes; it both mediated the effects of background characteristics and contributed uniquely to their determination. Finally, the importance of track placement for subjective outcomes was partly indirect. Track placement raised educational plans and self-esteem largely through its more immediate effects on the likelihood of associating with high status, high ability, and college-oriented peers and by either enhancing or depressing academic performance. Curriculum differentiation thus contributed to the maintenance of status advantages and disadvantages through the educational system by transmitting much of the influence of status origins upon a broad range of schooling outcomes.

These conclusions, while suggestive and potentially quite important, nevertheless are both incomplete and tentative. Their incompleteness refers to substantive issues not resolved by their analyses. Their tentativeness derives from data limitations, shared by all of the curriculum effects studies cited above, which might exaggerate the apparent importance of track placement.

With regard to data limitations, Alexander and McDill (1976) lacked precurriculum enrollment controls for their various school process outcomes. Under such circumstances, the interpretation of "curriculum effects" as such must be advanced with caution. Simply put, the various educational advantages evidenced by students enrolled in academic tracks may well have existed prior to their curriculum placement, and hence did not derive from such placement and attendant educational experiences.

Obviously, the "selection and alloca-

tion" interpretation of curriculum effects on plans, achievement, or any other educational outcome would be more secure were pre-enrollment levels of that outcome controlled. Lacking such longitudinal data, none of the curriculum effects research available to date has attempted such a stringent evaluation. This is one of the intentions of the present project. In so doing, we also will calculate the degree of upward bias in curriculum parameters estimated in the absence of such controls.

Assuming, as we do, that curriculum placement remains an important constraint on educational attainment even after more adequate evaluation, there still remain unresolved many important substantive issues. While much of the available literature has concentrated on the consequences of track membership, little has been learned of the mechanisms of curriculum sorting beyond the documentation of important social background and demographic differences in track placement. This neglect of allocative processes is particularly unfortunate in view of the apparent advantage of high status youth in achieving entry to college preparatory programs (Alexander and Eckland, 1975; Alexander and McDill, 1976). The present effort should help clarify the mechanisms by which high status origins actually promote college track placement. The direct transmission of status advantages will be distinguished from that deriving from status differences in socialization patterns and academic performance.

The present research also considers a more inclusive set of educational outcomes than have prior studies. Our models include as outcomes not only educational goals, academic achievement, and the social supports provided by significant others, but also whether the respondent has applied to college and, having applied, been accepted.

Goals of college attendance do not, in themselves, lead to college enrollment; indeed, no secondary school achievements nor background characteristics will secure a college education unless application for such is made. Knowledge of the common and unique determinants of both goals and goal-oriented behaviors will contribute to a better understanding of

why some youths successfully negotiate the transition from high school to college, while others, perhaps equally motivated, do not.

THEORETICAL AND ANALYTICAL MODEL

Figure 1 presents the model guiding the present analyses. The model is fully recursive between blocks; variables within blocks do not affect one another directly, but are permitted to have correlated structural disturbances. Ordinary least-squares regression procedures are used to estimate the structural parameters.

The specification of the model deserves brief comment. Socioeconomic background characteristics and the ascribed traits of race and sex are exogenous. Following common practice, ability is considered endogenous to these background characteristics. This specification permits estimation of the total effects of SES background and demographic characteristics on achievement, curriculum placement, and other outcomes through reduced-form equations and provides upperbound estimates of possible background influence on measured ability. This positioning of ability will not affect estimates of its total effects on the later outcomes.

A block of variables measured in the ninth grade, appearing next, intervenes between the exogenous and ability measures and the student's senior high school curriculum enrollment. Inclusion of these variables measuring previous academic achievement, motivations, and social supports will secure more accurate estimates of the unique contribution of curriculum enrollment to senior high school outcomes, and improve upon the model specifications employed in previous studies of curriculum effects. Additionally, we expect that these mechanisms will mediate much of the SES effects on curriculum enrollment, as well as increase our ability to predict enrollment in a college preparatory track.

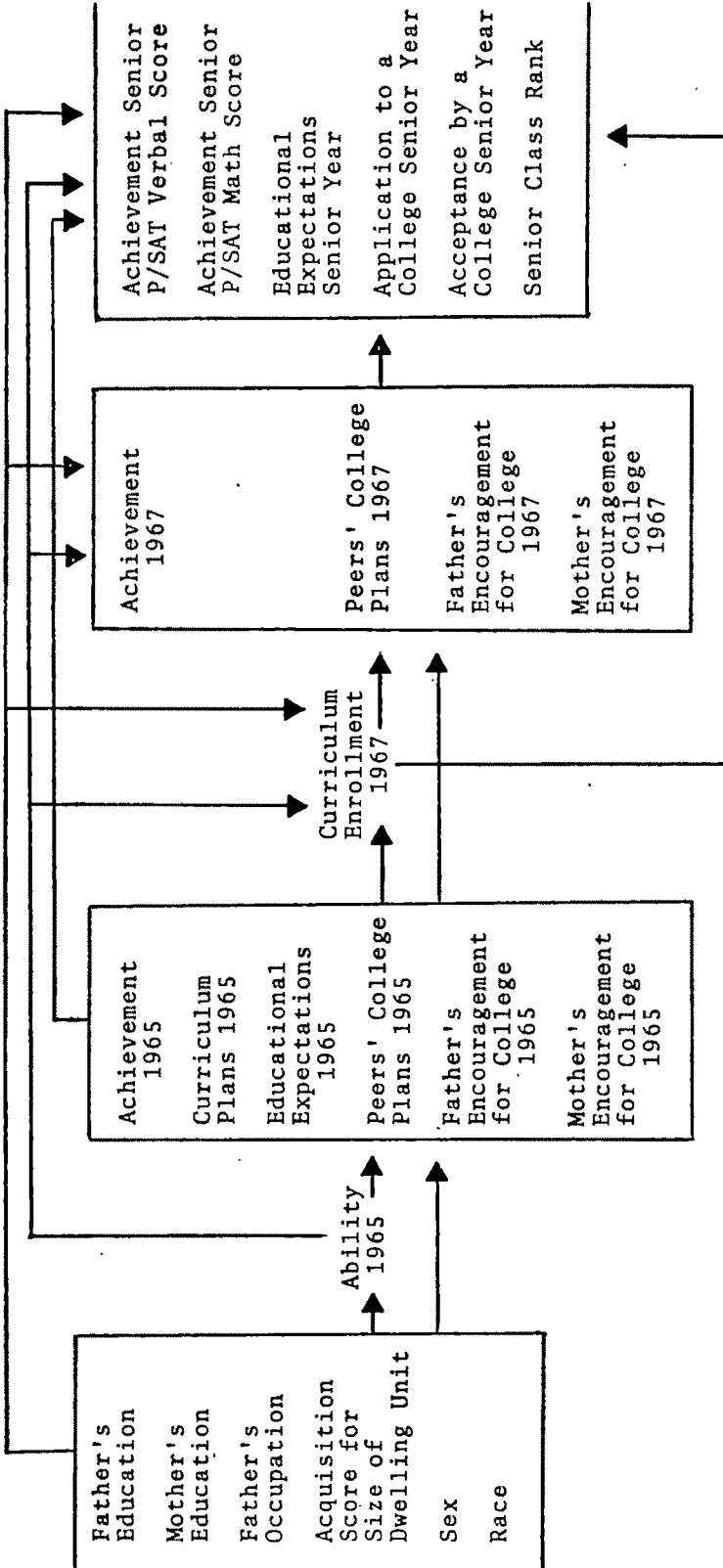
Senior high school curriculum enrollment, although ascertained in 1967, has been placed causally prior to other 1967 outcomes. The 1967 data were collected in

the eleventh grade, after most curriculum assignments had been made—usually in the tenth grade. We thus assume that curriculum placement may mediate influences of all prior variables upon all later outcomes.¹ Prior encouragements and achievements should influence track placement; curriculum membership should, in turn, affect later achievements as well as the likelihood of associating with college-oriented peers (Parsons, 1959). Finally, the labelling of students as either college-bound or not should temper the encouragement received from parents and others for pursuit of post-secondary education.

Standardized test scores, educational expectations, and senior class rank all have been examined in earlier research, and are included in this inquiry as well. The remaining two outcomes are behaviors directly related to the probability of post-secondary education. Application to college is generally necessary in order to attend college, and acceptance by an institution of higher education is a prerequisite. Table 1 documents the fact that a much larger proportion of seniors (63%) express college goals than have been accepted for college (19%), or even applied to a college (45%) as of January and February of their senior year. The extent of such discrepancies between expressed intentions and the behaviors necessary for their implementation suggests the importance of understanding the determinants of these goal-oriented behaviors.

¹ The grade structure of most schools in our sample (ten through twelve) also justifies this specification. Two of the eight high schools, however, have earlier entry levels, one in grade nine and one in grade seven, which possibly invalidates our assumption that ninth-grade measures are temporally, if not causally, prior to track placement. However, an analysis of covariance revealed few differences between these two schools and the other six. Knowledge of whether the student was enrolled in one of these two as opposed to any of the other six, together with all possible school interactions, added at most 2% explained variance over that obtained for the structural equations reported in the text. Also, our use of 1967 background data assumes that parental status characteristics are relatively stable over short periods of time (particularly during their children's adolescent years) and that later reports are likely more valid than earlier (1965) ones.

1967



* Variables are blocked for simplicity of presentation. The model is fully recursive; all variables directly affect all later ones. Variables within blocks are not assumed to be causally related; their structural disturbances are assumed to be correlated due to the omission of variables which would simultaneously affect all of them. See text for a brief justification of the temporal specification of the model, especially as concerning "curriculum placement."

Figure 1. Structural Model of Curriculum Placement Influences in the Adolescent Educational Attainment Process*

Table 1. Means, Standard Deviations, and Metric Information for Variables Employed in the Structural Model (N = 1607)

Variable		Overall Mean	Overall Standard Deviation	Within- School Standard Deviation*	Metric Information
•	Father's Education	12.75	2.75	2.51	Years of Schooling
•	Mother's Education	12.52	2.30	2.15	Years of Schooling
Back-	Father's Occupation	41.58	20.54	19.08	Duncan SEI
ground	Acquisition Index	9.62	2.72	2.62	Number of Rooms + Baths
•	Sex	.55	.50	.50	Female=1; Male=0
•	Race	.14	.35	.27	Black=1; Nonblack=0
1965	Ability	284.14	11.53	10.79	SCAT Test Score (Total)
•	Achievement 1965	1671.28	76.65	71.57	STEP Test Score (Total)
•	Curriculum Plans	.53	.50	.49	Academic=1; Other=0
•	Educational Plans	.69	.46	.45	College=1; Other=0
1965	Peers' College Plans	59.73	28.92	27.15	Percent Planning College
•	Father's Encouragement	83.11	14.25	13.96	Percent of Encouragement
•	Mother's Encouragement	83.89	13.77	13.54	Percent of Encouragement
1967	Curriculum Placement	.60	.49	.47	Academic=1; Other=0
•	Achievement 1967	1719.45	76.59	72.50	STEP Test Score (Total)
•	Peers' College Plans	60.70	26.88	25.68	Percent Planning College
1967	Father's Encouragement	82.50	14.35	14.05	Percent of Encouragement
•	Mother's Encouragement	82.83	14.01	13.69	Percent of Encouragement
•	Educational Plans	.63	.48	.45	College=1; Other=0
•	Applied to College	.45	.50	.49	Yes=1; No=0
Senior	Accepted by College	.19	.39	.37	Yes=1; No=0
Year	Achievement-Verbal	403.77	116.44	111.94	PSAT/SAT Test Score (Total)
•	Achievement-Math	416.25	129.24	124.13	PSAT/SAT Test Score (Total)
•	Senior Class Rank	62.83	29.57	29.56	Percent of Class Below

* These are the pooled within-school standard deviations employed in the analyses.

SAMPLE

The sample is a subset of the cohort of students in the Academic Growth Study (Educational Testing Service, Princeton, N. J.) who were contacted in the fall of 1965 while they were in the ninth grade, and recontacted in the fall of the eleventh (1967) and winter (January and February) of the twelfth grades in high school. In the ninth and eleventh grades, students were administered SCAT (School and College Ability Tests; see ETS, 1957a) verbal and quantitative batteries, six STEP tests (Sequential Tests of Educational Progress; see ETS, 1957b) and an attitude questionnaire. In the senior year the students completed a brief questionnaire largely concerned with post-secondary school plans, and were administered the PSAT (Preliminary Scholastic Aptitude Test) verbal and quantitative battery. The 1,607 students who form the sample for the present analyses are those who were enrolled in comprehensive (as opposed to vocational) high schools and for whom complete data were available on all variables of interest. They are enrolled in eight high schools in three large school districts across the United States; approximately 14% of the sample is black and approximately 55% is female.²

² Approximately 5,600 students in 27 high schools in 17 communities took the senior questionnaire in 1969; this figure establishes an upper bound on the longitudinal sample size. Information on respondents' race, however, was collected in only three communities in 1967. After elimination of students without race information and nonparticipants in any wave of data acquisition, we retained about 2,400 persons in ten high schools. Two schools were omitted because they were vocational rather than comprehensive schools; this left approximately 1,900 students. The exclusion of students due to missing data on any particular items in our analyses reduced the sample to approximately 1,600.

Previous research (e.g., Porter, 1974; DeBord et al., 1977) has documented race interactions in social-psychological achievement models similar to the one estimated here. An analysis of covariance revealed generally negligible race and sex interactions in these data, and thus permitted our use of a pooled sample with race and sex dummy variables. The increase in R^2 resulting from the inclusion of race and sex interaction terms to the various equations of our model was, for most outcomes, approximately 2% (absolute). Contrary to other studies of race and sex differences, we have used pooled within-school variances and covariances in our

VARIABLES

Background data on father's education and occupation, mother's education, and dwelling unit size were obtained from the 1965 and 1967 questionnaires. The 1967 responses were given priority due to minor coding irregularities in 1965; if 1967 responses were not available, however, 1965 responses were used.³ An ability measure—the total SCAT score in 1965—and sex and race information also are included. The latter two variables were ascertained from school records and from counselors or yearbooks, respectively.

Precurriculum enrollment (i.e., ninth grade) controls were obtained from the 1965 questionnaire. These included curriculum plans (dichotomized as academic/other),⁴ peers' educational plans (percent of friends planning to go to college, with values of 10, 30, 50, 70, 90% permissible), degree of maternal and paternal encouragement to go to college (with the five original Likert-type response options being assigned values corresponding to the percentages available for the peers' plans measure), and educational expectations (dichotomized as college/other). The total STEP battery score from the ninth grade was included as a measure of academic achievement.

Senior high school data on intervening

analysis (see below in text and also fn. 7). The proportion of blacks varied across the eight schools in our sample as follows: .085; .091; .058; .399; .000; .008; .924; .058. Thus, when purging our data of all school-to-school differences we, in effect, removed the marked between-school differences in racial composition. The remaining individual level within-school race effects thus are unconfounded by school-to-school differences in racial composition—a possibly quite important difference between our approach and that employed in other research on race and sex differences.

³ For the 1,607 cases the amount of data substitution for the four outcomes was: FAOCC 2.9%; MOED 1.6%; FAED 2.8%; ACQUIS .2%.

⁴ While the present project contrasts college preparatory enrollment with all other curricula, this crude classification may well mask important distinctions between nonacademic tracks. Future research on secondary school tracking should consider this possibility, although we believe the college/noncollege dichotomy reflects the major curriculum distinction insofar as issues of educational stratification are concerned.

and outcome variables were obtained during the junior and senior years. Peers' educational plans, and paternal and maternal encouragement for college were obtained from the 1967 eleventh-grade questionnaire. Curriculum enrollment was taken from school records when possible; where these data were unavailable, the student's report of his/her curriculum enrollment from the questionnaire was employed.⁵ The eleventh-grade total STEP score is employed as an academic achievement measure. Other outcome variables, obtained from the senior questionnaire, are the student's educational plans in the senior year and information on whether the student had applied to and, if so, had then been accepted by a college. The PSAT or SAT verbal and quantitative tests were administered by the ETS staff (PSAT) or obtained from school records (SAT). These testing data tap academic achievement at the end of high school. PSAT scores were multiplied by ten to place them on the same metric as the SAT, the latter scores being employed where available.⁶ Senior class rank, an indicator of relative achievement, was obtained from school records. The means, standard deviations and metric information for all variables are presented in Table 1.

ANALYSIS AND DISCUSSION

Our discussion will be organized around three general questions: (1) what are the mechanisms by which socioeconomic background affects curriculum placement; (2) what effect does curriculum placement have upon senior high school achievements (absolute and relative), goals, and behaviors designed to implement these goals when pre-enrollment controls are

included; (3) how severely biased are estimates of curriculum effects when pre-enrollment motivations and achievements are *not* controlled? Following precedent (Alexander and McDill, 1976; Heyns, 1974; Hauser et al., 1976), we conduct our analysis on a matrix of within-school variances and covariances. This procedure eliminates from the data matrix systematic differences from school to school, and thus focuses on selection and allocation processes *within* schools.⁷

Our results are presented in a series of tables organized around the major stages of the model depicted in Figure 1. Table 2 focuses on determinants of curriculum placement, with its terminal outcome being senior high school track membership. Table 3 extends the framework to consider other eleventh-grade outcomes, including academic achievement and social supports for college. Table 4 contains parameter estimates for the final stage of the model, senior year attainments. Throughout we employ the strategy of successive reduced-form equation estimation recommended by Alwin and Hauser (1975) for calculating the total, direct, and, through simple subtraction, indirect effects implicit in recursive models.

The SES-Curriculum Linkage

Only moderate proportions of variance in pre-enrollment motivations and social supports are accounted for by background characteristics and ninth-grade ability. Curriculum plans (24% explained vari-

⁵ The student's report was employed in 7.5% of the 1,607 cases.

⁶ Only about 14% of the scores on this variable are from the SAT. The data available to us are *either* PSAT or SAT scores, the former having been recorded if the school could furnish that information, and the latter being recorded for all other students from the ETS testing session in the winter of the senior year. To render the scores compatible, PSAT scores were multiplied by ten, an acceptable method for achieving this end (Julian Stanley, personal communication).

⁷ Allowing the slopes to vary across the eight schools increased the within-school explained variance an average of about 5% (absolute) per endogenous variable. The increment in R^2 was significant at .05, the level for 11 of the 18 outcomes in the structural model. In particular, the largest increments in explained variance were: 7% for EDEX-SR; 7% for APPLIED; 10% for ADMITTED; and 9% for SRRANK. The increments for senior year outcomes resulted from the addition of 126 interaction terms to each of the structural equations. Comparisons of the within-school parameter estimates across the eight schools revealed no substantively meaningful patterns, although the specific determinants of each senior year outcome sometimes differed markedly from school to school. Furthermore, when increments in explained variance were computed from R^2 's adjusted for the number of regressors in each equation the increments were essentially zero throughout.

Table 2. Parameter Estimates for the Reduced Form and Structural Equations in the Adolescent Attainment Model: Ability, Ninth-Grade Outcomes, and Curriculum Placement**

Dependent Variables	Independent Variables					
	FAED	MOED	FAOCC	ACQUIS	SEX	RACE
ABILITY	(.497) .116*	(.643) .128*	(.041) .072*	(-.227) -.055*	(-1.130) -.052*	(-9.864) -.251*
ACHV-9	(4.137) .145* (1.260) .044*	(4.267) .128* (.541) .016	(.202) .054* (-.034) -.009	(-1.735) -.064* (-.419) -.015	(4.860) .034 (11.406) .079*	(-50.807) -.195* (6.315) .024
CUPL-9	(.016) .084* (.007) .035	(.025) .110* (.012) .054*	(.004) .138* (.003) .107*	(.001) .004 (.005) .028	(-.078) -.079* (-.056) -.057*	(-.087) -.049* (.105) .059*
EDEX-9	(.017) .097* (.012) .065*	(.015) .074* (.008) .038	(.002) .081* (.001) .061*	(.002) .013 (.005) .029	(-.090) -.101* (-.077) -.086*	(.008) .005 (.122) .075*
PRPL-9	(1.274) .118* (.916) .085*	(1.234) .098* (.769) .061*	(.163) .115* (.134) .094*	(.165) .016 (.329) .032	(-6.065) -.111* (-5.250) -.096*	(-7.536) -.076* (-.420) -.004
FAEN-9	(.549) .099* (.447) .080*	(.159) .024 (.027) .004	(.066) .091* (.058) .080*	(.034) .006 (.081) .015	(-3.322) -.118* (-3.091) -.110*	(-.459) -.009 (1.558) .031
MOEN-9	(.093) .017 (.003) .001	(.449) .071* (.333) .053	(.054) .076* (.047) .066*	(.044) .009 (.085) .016	(-4.434) -.163* (-4.230) -.155*	(1.621) .033 (3.395) .069*
CURRIC	(.017) .089* (.007) .036 (.000) .002	(.028) .130* (.016) .071* (.011) .048*	(.003) .104* (.002) .071* (.001) .031	(.007) .037 (.011) .063* (.010) .053*	(-.143) -.151* (-.120) -.127* (-.100) -.106*	(-.088) -.051* (.109) .064* (.064) .037

ance) and peer plans (17% explained variance) are the most predictable of these early school process variables. In the structural equations, ability generally dwarfs the effects of socioeconomic factors and *uniquely* explains from 2 to 3% additional variance in the parental encouragement outcomes, 7 or 8% for educational goals and peers' plans, and 16% for curriculum plans.⁸ Females are

disadvantaged relative to males with regard to all outcomes except achievement and especially are lacking in parental and peer support for post-secondary schooling. Blacks, on the other hand, receive somewhat more maternal encouragement and have higher expectations (including plans for enrollment in a college track) once ability is controlled; these effects are, however, quite modest (see also Debord et al., 1977; Hout and Morgan, 1975; Kerckhoff and Campbell, 1977; Portes and Wilson, 1976).

In the reduced form equation, all SES and demographic variables (with the exception of the acquisition index) are significant determinants of track placement. Sex is the most important of these determinants, although the aggregated SES effect, at .263 (not reported in the tables), is larger. The inclusion of ability generally reduces but does not exhaust these direct

* Using a procedure devised by Heise (1972), we have calculated the aggregated effects of the four separate background indicators to facilitate comparison of SES influences with other variables in the model. The respective total and direct SES effects for the various precursriculum controls are: ACHV-9: .261, .049; CUPL-9: .262, .163; EDEX-9: .202, .139; PRPL-9: .265, .199; FAEN-9: .178, .144; MOEN-9: .132, .101. Thus, except for the three social support measures, the total effects of ability far exceed even the aggregated effects of the SES indicators and, throughout, ability accounts for much of the influence of status origins on later outcomes.

Table 2. Continued

Dependent Variables	Independent Variables							R ²
	ABILITY	ACHV-9	CUPL-9	EDEX-9	PRPL-9	FAEN-9	MOEN-9	
ABILITY								.154
ACHV-9								.128
	(5.791)							.772
	.873*							
CUPL-9								.084
	(.019)							.241
	.431*							
EDEX-9								.052
	(.011)							.117
	.277*							
PRPL-9								.098
	(.721)							.168
	.287*							
FAEN-9								.048
	(.204)							.069
	.158*							
MOEN-9								.044
	(.180)							.061
	.143*							
CURRIC								.102
	(.020)							.280
	.458*							
	(.007)	(.001)	(.204)	(.111)	(.001)	(.002)	(-.000)	.380
	.151*	.179*	.212*	.106*	.069*	.073*	-.002	

* Coefficient greater than or equal to 1.96 times its standard error; standardized and raw (in parentheses) coefficients.

** FAED, Father's Education; MOED, Mother's Education; FAOCC, Father's Occupation; ACQUIS, Acquisition Index; ACHV-9, STEP Achievement in Ninth Grade; CUPL-9, Curriculum Plans in Ninth Grade; EDEX-9, Educational Expectations in Ninth Grade; PRPL-9, Peers' College Plans in Ninth Grade; FAEN-9, MOEN-9, Father's and Mother's Encouragement for College in Ninth Grade; CURRIC, Curriculum Placement; ACHV-11, STEP Achievement in Eleventh Grade; PRPL-11, Peers' College Plans in Eleventh Grade; FAEN-11, MOEN-11, Father's and Mother's Encouragement for College in Eleventh Grade; P/SAT-V, P/SAT-M, Verbal and Math Achievement Scores in Twelfth Grade; SRRANK, Senior Class Rank; EDEX-SR, Senior Educational Expectations; APPLIED, Application to College; ADMITTED, Acceptance by a College.

effects. The blocked SES effect, for example, drops to .163 after measured ability is added. When controls for junior high school motivations and social supports are added to the equation, these exogenous influences are reduced still further, generally to the point of triviality. Even the aggregated SES effect is quite small, at .091. Thus, almost two-thirds of the influence of socioeconomic origins on track placement is accounted for by the various precurriculum school process variables considered here—academic ability and achievement, educational goals

and curriculum plans and the social supports provided by parents and friends. The only background characteristic whose importance is largely independent of these intervening mechanisms is sex, whose structural coefficient remains two-thirds the size of its reduced form counterpart (Alexander and Eckland, 1974, similarly found little transmission of sex differences through such mechanisms). Even the substantial influence of ninth-grade ability on eleventh-grade curriculum placement is largely mediated through these intervening mechanisms, with total and direct ef-

Table 3. Parameter Estimates for the Reduced Form and Structural Equations in the Adolescent Attainment Model: Eleventh-Grade Outcomes**

Dependent Variables	Independent Variables						
	FAED	MOED	FAOCC	ACQUIS	SEX	RACE	ABILITY
ACHV-11	(.316)	(.956)	(.011)	(.212)	(-6.346)	(-13.022)	(1.525)
	.011	.028*	.003	.008	-.044*	-.049*	.227*
	(.309)	(.751)	(-.004)	(.026)	(-4.393)	(-14.268)	(1.397)
	.011	.022	-.001	.001	-.030*	-.054*	.208*
PRPL-11	(.248)	(1.032)	(.046)	(.404)	(-5.507)	(1.098)	(.025)
	.024	.086*	.034	.041	-.107*	.012	.010
	(.244)	(.932)	(.039)	(.314)	(-4.555)	(.491)	(-.038)
	.024	.078*	.029	.032	-.088*	.005	-.016
FAEN-11	(.372)	(.060)	(.022)	(.164)	(-1.742)	(1.019)	(-.026)
	.066*	.009	.030	.031	-.062*	.020	-.020
	(.370)	(.002)	(.018)	(.113)	(-1.201)	(.674)	(-.062)
	.066*	.000	.024	.021	-.042	.013	-.047
MOEN-11	(.207)	(.291)	(.015)	(.104)	(-3.119)	(2.625)	(-.048)
	.038	.046	.020	.020	-.113*	.053*	-.038
	(.205)	(.227)	(.010)	(.046)	(-2.506)	(2.234)	(-.088)
	.038	.036	.014	.009	-.091*	.045*	-.069

fects of .458 and .151, respectively. Moreover, these pre-enrollment school process influences contribute uniquely to the prediction of track membership, as evidenced by the increment in R^2 from .280 to .380 upon their addition to the curriculum equation. In the final (structural) equation, the major direct determinants of enrollment in a college track are ability, achievement, and ninth-grade curriculum plans. In contrast, the *direct* effects of all status background indicators and race are negligible.

These results suggest that we have indeed identified some of the important linkages by which high status origins enhance one's prospects for enrollment in a college track (or, conversely, by which low status origins impede such prospects). High status youths benefit from a succession of modest advantages over the course of their early school careers. These, in the aggregate, practically exhaust the relevance of status origins for curriculum placement. By the time of entry into secondary school, higher status students already exceed lower status youths in plans for college and plans to pursue an academic program of study.⁹ They also are

more involved in peer networks supportive of academic pursuits and receive more parental support for college plans than their lower status, even equally able lower status, classmates. These, then, are the proximate determinants of track placement and high status youths are somewhat advantaged on each of them. Other ascribed statuses, on the other hand, affect track placement quite differently. The primary reason for the lesser likelihood of blacks enrolling in a college track is their lower average test scores. With ability controlled, blacks actually are somewhat *more likely* than whites to enroll in an academic program. Thomas et al. (1977) report a similar finding for a larger, nationally representative sample of high school students. Finally, the disadvantage experienced by women with regard to the likelihood of enrolling in a college track is largely independent of all of these mechanisms.

Curriculum Effects on Eleventh-Grade Outcomes

Table 3 presents the results for eleventh-grade outcomes. In Table 3 (and

⁹ The notable effect of ninth-grade curriculum plans on subsequent enrollment suggests that there may be considerable voluntarism in tracking decisions. Thus, the contrast between selection-allocation perspectives and socialization models

(Kerckhoff, 1976) probably should not be too sharply drawn until more evidence is available on the bases of selection. Nevertheless, the consequences of decisions, once made, might be allocative in nature, regardless of the mechanics of the selection process.

Table 3. Continued

Independent Variables							R ²
ACHV-9	CUPL-9	EDEX-9	PRPL-9	FAEN-9	MOEN-9	CURRIC	
(.625)	(8.037)	(.301)	(.061)	(-.008)	(-.031)		.781
.617*	.054*	.002	.023	-.002	-.006		
(.602)	(4.047)	(-1.871)	(.038)	(-.057)	(-.030)	(19.512)	.790
.595*	.027	-.012	.014	-.011	-.006	.126*	
(.075)	(1.767)	(3.644)	(.255)	(.108)	(.044)		.312
.208*	.034	.063*	.270*	.059	.023		
(.063)	(-.177)	(2.586)	(.244)	(.085)	(.045)	(9.507)	.331
.177*	-.003	.045	.258*	.046	.024	.174*	
(.009)	(1.764)	(2.178)	(-.016)	(.354)	(.059)		.240
.045	.061*	.069*	-.030	.352*	.057		
(.003)	(.658)	(1.576)	(-.022)	(.341)	(.060)	(5.409)	.260
.013	.023	.050	-.042	.339*	.058	.181*	
(.010)	(1.968)	(2.887)	(.009)	(.088)	(.297)		.260
.052	.070*	.094*	.017	.090*	.294*		
(.003)	(.714)	(2.204)	(.001)	(.073)	(.298)	(6.132)	.288
.015	.025	.072*	.003	.074*	.295*	.210*	

* Coefficient greater than or equal to 1.96 times its standard error; standardized and raw (in parentheses) coefficients.

** See Table 2 for variable abbreviations.

Table 4) we are concerned mainly with establishing the extent to which curriculum placement conditions access to various educational resources and either enhances or retards academic achievements. To simplify the presentation somewhat, we disregard the causal relations among precurriculum variables and treat them all as predetermined relative to curriculum placement and subsequent outcomes.

The curriculum effects in Table 3 are pervasive, but modest throughout. College track placement enhances eleventh-grade achievement even when ability and ninth-grade achievement are controlled. The gain on the combined STEP score is about one-fourth (28%) of a within-school standard deviation. Other influences on achievement are largely as anticipated, with measured ability and prior achievement having the only other appreciable effects (even the aggregated total effect of SES background is trivial, at .038).

Each of the structural equations for parental and peer support for post-secondary schooling exhibits a similar pattern of influences. Having had college-oriented friends in the ninth grade is the principal determinant of the corresponding eleventh-grade associations, while curriculum and prior achievements have secondary, and about equal, implications

for such peer relations. Parents' encouragement of their children's aspirations for college also is responsive to track placement (or, at least, students in academic tracks are likely to perceive such encouragement). In the structural equations for parental encouragement, curriculum placement has the second strongest effect (second only to the corresponding ninth-grade parental encouragements). All other *direct* effects in the three equations are negligible. Even the blocked SES effects range from only .074 to .116. For all interpersonal outcomes, the importance of curriculum placement is largely in transmitting prior influences. But, it also *uniquely* induces an additional 2 to 3% explained variance. These strike us as rather impressive figures in view of the extensive controls on prior achievements, abilities, and motivations in these equations.

Senior Year Outcomes

Table 4 presents the last portion of the model, which involves a broad range of senior-year outcomes—verbal and quantitative standardized test scores, senior class rank, educational goals, whether the student had applied to college and whether s/he had been admitted.

Curriculum placement contributes little

Table 4. Parameter Estimates for the Reduced Form and Structural Equations in the Adolescent Attainment Model: Twelfth-Grade Outcomes**

Dependent Variables	Independent Variables								
	FAED	MOED	FAOCC	ACQUIS	SEX	RACE	ABILITY	ACHV-9	CUPL-9
P/SAT-V	(1.577)	(1.144)	(.264)	(-.911)	(-5.433)	(11.991)	(2.781)	(.831)	(10.995)
	.035	.022	.045*	-.021	-.024	.029	.268*	.531*	.048*
	(1.571)	(.973)	(.251)	(-1.065)	(-3.810)	(10.955)	(2.675)	(.812)	(7.679)
	.035	.019	.043*	-.025	-.017	.027	.258*	.519*	.033
	(1.282)	(.462)	(.249)	(-1.109)	(-3.392)	(19.131)	(1.845)	(.441)	(4.946)
P/SAT-M	.029	.009	.042*	-.026	-.002	.047*	.178*	.282*	.022
	(1.429)	(.639)	(-.029)	(1.288)	(-47.900)	(-18.284)	(4.442)	(.494)	(20.775)
	.029	.011	-.004	-.027	-.192*	-.040*	.386*	.285*	.082*
	(1.411)	(.145)	(-.065)	(.841)	(-43.207)	(-21.276)	(4.135)	(.439)	(11.190)
	.029	.003	-.010	.018	-.173*	-.047*	.359*	.253*	.044*
SRRANK	(1.155)	(-.415)	(-.071)	(.776)	(-39.589)	(-15.158)	(3.485)	(.135)	(8.953)
	.023	-.007	-.011	.016	-.158*	-.034*	.303*	.078*	.035
	(.334)	(.185)	(-.032)	(-.313)	(4.309)	(-2.893)	(.412)	(.090)	(-.495)
	.028	.013	-.021	-.028	.072*	-.027	.150*	.217*	-.008
	(.333)	(.148)	(-.035)	(-.346)	(4.658)	(-3.116)	(.389)	(.085)	(-1.208)
EDEX-SR	.028	.011	-.022	-.031	.078*	-.029	.142*	.207*	-.020
	(.265)	(.010)	(-.038)	(-.380)	(5.370)	(-1.246)	(.220)	(.008)	(-1.659)
	.023	.001	-.024	-.034	.090*	-.012	.080	.020	-.027
	(.009)	(.005)	(.001)	(-.001)	(-.083)	(.176)	(.005)	(.001)	(.087)
	.051	.026	.060*	-.007	-.091*	.106*	.129*	.090*	.094*
APPLIED	(.009)	(.002)	(.001)	(-.004)	(-.048)	(.153)	(.003)	(.000)	(.016)
	.050*	.009	.049*	-.026	-.053*	.093*	.074	.026	.017
	(.007)	(-.001)	(.001)	(-.005)	(-.030)	(.158)	(.002)	(-.000)	(.012)
	.040	-.004	.043	-.032	-.033	.096*	.059	-.062	.013
	(.019)	(.004)	(.001)	(.007)	(-.028)	(.116)	(.006)	(.001)	(.112)
ADMITTED	.098*	.017	.030	.038	-.029	.065*	.125*	.143*	.112*
	(.019)	(.001)	(.001)	(.005)	(-.001)	(.099)	(.004)	(.001)	(.057)
	.097*	.005	.022	.025	-.001	.056*	.086	.097*	.058*
	(.017)	(-.001)	(.000)	(.004)	(.014)	(.107)	(.003)	(.000)	(.053)
	.089*	-.005	.017	.020	.014	.061*	.067	.009	.053*
ADMITTED	(.001)	(.001)	(.002)	(.002)	(.001)	(.025)	(.002)	(.001)	(.043)
	.008	.003	.086*	.013	.001	.018	.048	.140*	.057
	(.001)	(-.001)	(.002)	(.000)	(.016)	(.015)	(.001)	(.001)	(.012)
	.007	-.006	.080*	.003	.022	.011	.019	.106*	.016
	(.000)	(-.002)	(.001)	(-.000)	(.025)	(.014)	(.001)	(.000)	(.010)
	.000	-.013	.076*	-.001	.033	.010	.019	.077	.013

* Coefficient greater than or equal to 1.96 times its standard error; standardized and raw (in parentheses) coefficients.

** See Table 2 for variable abbreviations.

Table 4. Continued

Dependent Variables	Independent Variables									R ²
	EDEX-9	PRPL-9	FAEN-9	MOEN-9	CURRIC	ACHV-11	PRPL-11	FAEN-11	MOEN-11	
P/SAT-V	(-5.126)	(.117)	(-.208)	(.037)						.672
	-.020	.028	-.026	.005						.675
	(-6.932)	(.097)	(-.248)	(.039)	(16.216)					.710
	-.028	.024	-.031	.005	.068					.601
P/SAT-M	(-6.494)	(.076)	(-.279)	(-.015)	(2.112)					.621
	-.026	.019	-.035	-.002	.009	(.613)	(.002)	(.149)	(.213)	.640
	(12.110)	(.252)	(-.099)	(-.386)		.397*	.000	.019	.026	.163
	.044*	.055*	-.011	-.042						.165
SRRANK	(6.890)	(.196)	(-.215)	(-.382)	(46.876)					.189
	.025	.043*	-.024	-.042	.177*					.298
	(6.689)	(.149)	(-.224)	(-.485)	(33.905)					.378
	.024	.033	-.025	-.053*	.128*	(.492)	(.114)	(-.001)	(.376)	.399
EDEX-SR	(3.400)	(.035)	(-.011)	(-.001)		.287*	.024	-.000	.041	.283
	.051	.032	-.005	-.000						.324
	(3.012)	(.031)	(-.019)	(-.001)	(3.486)					.339
	.045	.029	-.009	-.000	.055					.132
APPLIED	(3.139)	(.009)	(-.043)	(.028)	(.590)					.155
	.047	.009	-.020	.013	.009	(.120)	(.079)	(.094)	(-.116)	.298
	(.214)	(.002)	(.002)	(-.000)		.295*	.069*	.044	-.054	.378
	.211*	.112*	.066*	-.009						.399
ADMITTED	(.176)	(.001)	(.001)	(-.000)	(.346)					.283
	.173*	.087*	.039	-.008	.360*					.324
	(.166)	(.001)	(.000)	(-.001)	(.295)	(.001)	(.002)	(.002)	(.000)	.339
	.163*	.057*	.011	-.024	.306*					.132
	(.146)	(.002)	(.001)	(-.000)		.110*	.121*	.064*	.031	.155
	.134*	.125*	.025	-.003						.164
	(.117)	(.002)	(.000)	(-.000)	(.265)					.324
	.107*	.107*	.006	-.002	.256*					.339
	(.109)	(.001)	(-.001)	(-.000)	(.218)					.132
	.101*	.083*	-.020	-.009	.211*	(.001)	(.002)	(.002)	(.000)	.155
	(.074)	(.001)	(-.000)	(.001)		.119*	.096*	.065*	.006	.164
	.089*	.094*	-.004	.035						.023
	(.057)	(.001)	(-.000)	(.001)	(.154)					.164
	.069*	.081*	-.018	.035	.194*					.164
	(.051)	(.001)	(-.001)	(.001)	(.130)	(.000)	(.001)	(.002)	(.001)	.164
	.061*	.065*	-.042	.023	.164*	.025	.072*	.058	.023	.164

* Coefficient greater than or equal to 1.96 times its standard error; standardized and raw (in parentheses) coefficients.

** See Table 2 for variable abbreviations.

to *verbal* achievement in the senior year, either before or after the inclusion of eleventh-grade outcomes. In the structural equation for the verbal P/SAT, eleventh-grade achievement, ninth-grade achievement, and ability are the major direct determinants, in that order. In the structural equation for math performance, on the other hand, track placement is of some importance in addition to prior achievements and aptitudes. Being in a college track is worth about 26% of an overall and 27% of a pooled within-school standard deviation on the P/SAT quantitative battery. These differences in the determinants of verbal and quantitative performance might suggest that math achievement benefits more than verbal performance from specialized course enrollment.¹⁰

Although college preparatory work generally enhances students' absolute achievements during senior high school, it has virtually no effect on a relative measure—class rank. These data suggest that higher grades, and thus higher class rank, are not disproportionately allocated to students in academic tracks over and above what might be expected from their somewhat higher abilities and achievements vis-à-vis noncollege preparatory track students. The distribution of grades within each broad curriculum thus appears to allow excellence as well as failure to be accorded for mastery, or lack thereof, of program-specific subject matter. These findings, somewhat at variance with previous results (Alexander and McDill, 1976), suggest the need for detailed data on the administrative practices that govern relative achievement. Such practices might include differential weighting of grades in college versus nonacademic tracks, extra credit for accelerated or advanced placement courses, and the like (see fn. 7). Finally, females, on the average, rank about five percentage points above men. Obviously, we cannot

ascertain from these data whether "femaleness" in this context reflects better study habits, more diligence regarding completion of homework assignments, less disruptive classroom behavior—in short, closer approximation to the ideal student prototype once ability and achievement levels are controlled—or whether it represents only the teacher's presumption that young women try harder and are more cooperative than males (see Boocock, 1972, for a discussion of this issue).

Track placement appears to have especially marked consequences for goal orientations. Knowledge of a student's curriculum membership uniquely explains about 8% of the variance in senior year educational goals even after controlling for prior (ninth-grade) goals. In fact, its structural effect is just under twice that of these earlier expectations and its addition to the prediction equation reduces to triviality the direct impact of exogenous variables (other than race), ability, ninth-grade achievement and curriculum goals, and earlier social supports for college attendance. When later (eleventh-grade) controls are added, the direct effect of curriculum is reduced only slightly, and remains by far the most important direct determinant of educational goals. Enrollment in a college preparatory track increases by about 30% the probability that students will plan in their senior year to continue their education in comparison to equally able, motivated, and encouraged youth in nonacademic programs. Prior goals, eleventh-grade achievements, status origins, and peer supports also impact moderately on expectations.

In view of the extensive controls in the structural equation, this direct impact of curriculum on education goals is, to us, especially noteworthy. However, interpretation of these effects is ambiguous. Many students in college-preparatory tracks may express college intentions merely because they have been chosen and labelled by their high schools as having college potential. This "halo" need not motivate behaviors designed to secure further education, and thus may be rather unimportant in transmitting status advan-

¹⁰ There is, in fact, some suggestive evidence regarding the greater responsiveness of mathematics achievement, as opposed to verbal achievement, to specialized curricula and coaching. See, for example, College Entrance Examination Board (1968); McDill and Rigsby (1973: 63–5); Shaycoft (1967).

tages to these individuals as adults. In other instances college orientations might reflect the fact that students in academic programs have been disproportionately counselled and encouraged to apply to college by teachers, counsellors, and parents and as a result have done so (Heyns, 1974). Thus, the fact that college track membership encourages motivating aspirations is an important product of curriculum differentiation and allocation. In still other circumstances, these educational plans may only reflect certain knowledge of prior acceptance by a college, and thus may have little motivational relevance (Kerckhoff, 1976). Curriculum allocation in this instance might merely affect the initial mechanics of college application, rather than structure ambition. In short, interpretation of the impact of track segregation upon educational goals would differ decidedly depending upon which of these various alternatives actually obtained.

In the prediction of application to college, the addition of track membership to the equation containing background and pre-enrollment controls increases the coefficient of determination 4%. Curriculum is identified as the major direct determinant of the likelihood of so applying, substantially reducing the direct contributions of ability and junior high school achievements. As before with educational goals, the inclusion of controls for eleventh-grade supports and achievements reduces only slightly the salience of an academic curriculum for college application. Realizing the importance of the application procedure for continuing one's education beyond high school, these results are quite relevant to adult attainment and indicate the fruitfulness of a structural view of adolescent attainment. Enrollment in a college-preparatory program markedly increases the likelihood (by about 22% as suggested by these data) of applying to college over what would be expected for students similar in all other respects save their enrollment in a nonacademic track.

Within the confines of our model, the role of academic track membership in securing admission to a college is similar to its role in formulating college orienta-

tions and attempts to secure college acceptance. It uniquely contributes to explained variation (2%) and mediates prior influences. Again, it remains clearly the most important direct determinant of the senior year outcome—in the structural equation in which controls are included for background, ability, and both pre- and postenrollment achievements and social supports, membership in an academic program in high school increases by .13 the probability of being accepted by a college. Other direct influences, although statistically significant, are substantively trivial. This holds even for the aggregated effects of SES, with total and direct effects of .094 and .073, respectively.

Expectations to attend, application to, and acceptance by a college are by no means equivalent. As noted earlier, 63% of these students express college goals, 43% have applied to and only 19% have been accepted by a college. A comparison of the raw coefficients for the impact of curriculum placement in the structural equations for these three outcomes reveals an inverse relationship between the impact of academic track membership and the concreteness of the outcome. Plans to attend college, perhaps a totally ungrounded aspiration, are enhanced most by academic track placement; the taking of steps to secure admission to a college next so; and actual acceptance by a college least. The overall predictability of these outcomes from our model also is related inversely to the probability of attending college. We can predict 40% of the variance for college plans, 34% for application to and 16% for acceptance by a college.

It should also be noted, however, that estimates of the direct effects of curriculum presented here—certainly in the cases of ADMITTED and possibly in the case of APPLIED—likely are somewhat upwardly biased due to misspecification of the relationships among these endogenous variables. The causal relationships between goals, application, and acceptance are not immediately apparent due to their simultaneous measurement late in the senior year of high school. The exact nature of plans for further education are, in this model (as in most such research), un-

determined. They might reflect vague aspirations, motivating influences, or concrete knowledge of prior acceptance by college. In short, depending upon the amount of information available to students regarding their actual likelihood of attendance, causality potentially will vary across different subsamples of students. Further effort to disentangle the relationships among these three outcomes would be tangential to our major interests and will not be pursued here, although an effort to do so is in progress (Cook and Alexander, 1977).

The Importance of Pecurriculum Controls

The pre-enrollment (i.e., junior high school) controls are generally, with the exception of ninth-grade achievement, rather poorly predicted by our model. This suggests that extensive and perhaps novel elaboration of the traditional school process framework would be required to do better. Nevertheless, these outcomes themselves do contribute to our understanding of how curriculum placement affects adolescent achievement. Pecurriculum enrollment controls for achievement, motivations, and social supports generally serve several functions. As noted earlier, they mediate almost all of the effects of exogenous SES variables on curriculum enrollment, as well as absolutely increase by 10% the explained variance in enrollment in a college track. They also mediate prior influences and, more importantly, induce substantial unique variability in senior high school outcomes subsequent to curriculum enrollment relative to that obtained when background and ability alone are used to explain these outcomes. A further reason for concerning ourselves with junior high school outcomes is documented in Table 5: when these controls are *not* included in the prediction equation for senior high school outcomes, the total effects of curriculum enrollment are markedly overestimated as compared to the criterion estimates in the correctly specified model (see Table 5: II. A-C). These pre-enrollment controls generally attenuate estimates of curriculum influence from 20 to 50%. While such re-

ductions are considerable and suggest the need for caution in accepting uncritically the results of similarly misspecified models, this quite stringent evaluation still reveals the consequences of track placement to be pervasive and frequently substantial. Thus, while earlier inquiries may have exaggerated the precise magnitude of tracking influences, their interest in selective and allocative mechanisms as constraints on adolescent achievement has not been misplaced.

SUMMARY AND DISCUSSION

Our analyses address many of the issues raised but not resolved in earlier research on high school tracking. Our results suggest that the socioeconomic characteristics of students do affect their curriculum enrollment but do so almost exclusively through their influence upon achievements, goals, and encouragements by others in junior high school. Although these junior high school outcomes are poorly predicted from the background characteristics included in our model, they nevertheless are of considerable consequence. They mediate prior influences, are quite important for senior high school curriculum placement, and predict later encouragements and goals better than any other variables included in the analysis.

Even with pre-enrollment controls, however, the importance of curriculum placement for junior-year and senior-year outcomes is marked. In particular, tracking consistently affects educational goals, achievements, and goal-oriented behaviors in the twelfth grade, and is often the most important factor of those included in our model. Being in a college track increases the probability of applying to college and enhances one's prospects for being admitted. In these ways, sorting processes within high school may substantially affect later socioeconomic attainments.

Concerning the proper role of curriculum differentiation in high school education, two counterposed perspectives may be identified in the sociological and educational literatures and in popular thought. The first maintains that resources should be allocated where they can

Table 5. Estimates of Bias in Curriculum Effect Parameters Due to Omission of Pre-Enrollment Controls (N=1607)

Coefficient for Curriculum Enrollment**	Dependent Variables									
	ACHV-11	PRPL-11	FAEN-11	MOEN-11	EDEX-SR	APPLIED	ADMITTED	P/SAT-V	P/SAT-M	SRRANK
A. Total Effects of Curriculum, NO Pre-Enrollment Controls	.1764	.2662	.2790	.3123	.4411	.3346	.2461	.1091	.2129	.0819
B. Total Effects of Curriculum WITH Pre-Enrollment Controls	.1264	.1739	.1808	.2104	.3595	.2561	.1943	.0680*	.1774	.0554*
C. Total Effects Bias	39.6%	53.1%	54.3%	48.4%	22.7%	30.7%	26.7%	20.0%
D. Structural Effect of Curriculum, NO Pre-Enrollment Controls					.3474	.2505	.1818	.0047*	.1398	.0110*
E. Structural Effect of Curriculum WITH Pre-Enrollment Controls					.3063	.2114	.1636	.0089*	.1283	.0094*
F. Structural Effects Bias					13.4%	18.5%	11.1%	9.0%
G. Structural R ² :										
NO Controls	.7108	.2480	.1207	.1528	.3724	.3143	.1553	.6955	.6347	.1875
WITH Controls	.7905	.3310	.2604	.2879	.3992	.3388	.1642	.7095	.6404	.1894

* Not significantly different from zero at $\alpha \leq .05$.

** Standardized coefficients reported.

achieve maximum returns (Parsons, 1959). More able students who have, in elementary and junior high school, demonstrated high achievement levels and are motivated to pursue higher education should be provided access to an academically oriented, enriched high school learning environment. Students of lesser ability and past performance, often despite desires for a college education, are properly channeled into general or vocational tracks. Thus, the argument goes, each group of students can be taught at a level appropriate to its potential. College-bound students are challenged, stretched (Rosenbaum, 1975), and generally encouraged to achieve to their potential (Cicourel and Kitsuse, 1963). They benefit from (1) the more advanced material to which they are exposed, (2) the faster pace at which they can progress by being insulated from their less able peers, (3) the comforting knowledge that they are, indeed, college material, and (4) the atmosphere of encouragement from equally able peers, interested teachers and counsellors (Heyns, 1974). Students who are judged, according to ability, past performance, interests and motivation, to be incapable of performing college work likewise benefit from their assignment to a noncollege program of study with their equals: (1) material is presented to them which is relevant to their likely future adult circumstances; (2) their self-concept does not suffer from failure in competition with much more academically able students (see Coleman et al., 1966); and (3) material is presented at a pace consistent with their ability to absorb it. Curriculum differentiation therefore allocates scarce resources in a manner beneficial to all concerned, including the collective welfare of the larger social system.

Others, critics of tracking as it is presently practiced, suggest that tracking channels scarce resources to those who have the least need for them. Students in noncollege tracks are denied access to students, teachers, counselors, and information which would broaden their interests, challenge their abilities (Rosenbaum, 1975), and improve their performance (Heyns, 1974; Cicourel and Kitsuse, 1963). They are discouraged from compe-

tition with initially more advantaged students and hence are not required or even encouraged to strive for academic excellence. They are looked down upon by persons in academic tracks as being somewhat stupid, suffer from feelings of inferiority, and fail to develop attitudes and insights concerning education and institutional functioning which would allow them to compete successfully with their more advantaged classmates for post-schooling resources and rewards (Gintis, 1971; Bowles and Gintis, 1976). Moreover, they are shunted into curricula which will impede their prospects for success in college should they persevere in their college aspirations (Ramsøy, 1965) and more likely will be relegated to junior and community colleges, further "cooling out" their ambition (Clark, 1960; Karabel, 1972). They are, thereby, while still adolescents, subjected to social forces beyond their control, or at least whose implications they cannot fully appreciate—which may limit in important ways their prospects for adult success. Thus, curriculum differentiation benefits the advantaged and discriminates against those most in need of additional resources; it especially serves the interests of higher status parents who exploit such mechanisms to legitimate and perpetuate their children's success. As the above suggests, the literature critical of current school organization typically develops two closely linked, but separable, themes. One emphasizes socioeconomic biases in educational policies and practices; the second involves the preference given intellectually advantaged, rather than deprived, youth.

Our data actually are somewhat supportive of the positions taken by both proponents and critics of tracking. The major determinants of curriculum assignment are ability, junior high school achievement, and curriculum and educational aspirations in the ninth grade. However, over 60% of the variance in placement is left unexplained by these factors; thus, criteria other than objective ability and performance are quite important in the allocation of resources to students. Furthermore, the total effects of the background variables in our model on curriculum placement document the ten-

dency for higher status students to be streamed disproportionately into college preparatory curricula and hence to receive its attendant benefits.

Placement in a college track *does* enhance achievements, goals, and social supports in senior high school and markedly increases the probability of application to and acceptance by a college. Thus, the advantages accruing to such students are cumulative. Their favored backgrounds and early academic achievements increase the likelihood of enrollment in a college track, which accrues additional, wide-ranging educational benefits. Conversely, students in nonacademic tracks are, by virtue of such assignment, substantially disadvantaged in their future educational careers. While such students are recognized by both proponents and critics of tracking as being initially academically disadvantaged, the former think it most judicious that academic resources be directed elsewhere while the latter contend that efforts to close, rather than widen, the gap between high and low achievers should be given first priority (Rosenbaum, 1975).

If students were in fact assigned to curricula *strictly* on the basis of either merit (i.e., performance and ability) or ascriptive statuses (i.e., socioeconomic status, race, sex), then the consequences of these tracking influences would at least conform to one or the other of the scenarios outlined above. Actually, each is only in part correct, and the situation is even more complex than either perspective suggests. Clearly, the different curricula do independently affect the achievements of their members. They improve those of academic students and depress those of general and vocational track members. Under an efficient meritocracy, however, students of equivalent ability and performance would never be assigned to different curricula. Hence, no student's objective potential would be stifled and preference for such practices would depend on whether one thought that resources are best allocated to the most or to the least needy. However, there actually is considerable slippage in the process of curriculum sorting. Factors entirely unrelated to objectively assessed performance and

potential are important determinants of track assignment, and, hence, markedly influence a student's subsequent academic career through the simple administrative act of track placement. Two students of equal ability, motivation and past performance can be, and often are, assigned to different curricula!

These data cannot resolve what is fundamentally a political and ideological question: whether benefiting the advantaged or the disadvantaged is the more proper use of public resources. It is clear, however, that achievement-related criteria are not the sole, or even major, bases of curriculum assignment. Students of similar potential often are placed, for as yet undetermined reasons, in different tracks. In such cases, curriculum assignment expands opportunities for one group and constricts them for the other. At the same time, *direct* socioeconomic ascription in track placement is almost negligible. Thus, to a considerable degree the benefits associated with enrollment in a college program are available *entirely independent* of status origins and academic achievement. This suggests that neither functional nor conflict theories adequately characterize the role of curriculum differentiation in educational and social stratification.

Differential tracking in secondary schools thus introduces academic inequalities where none previously existed, and in so doing contributes independently to educational and socioeconomic inequalities. Such unrecognized consequences of administrative practices might be thought undesirable and undeserved by proponents of each of the above perspectives on the social organization of schooling.

REFERENCES

- Alexander, K. L. and B. K. Eckland
1974 "Sex differences in the educational attainment process." *American Sociological Review* 39:668-81.
- 1975 "School experience and status attainment." Pp. 171-210 in S. D. Dragastin and G. H. Elder, Jr. (eds.), *Adolescence in the Life Cycle: Psychological Change and Social Context*. Washington, D.C.: Hemisphere.

- Alexander, K. L. and E. L. McDill
1976 "Selection and allocation within schools: some causes and consequences of curriculum placement." *American Sociological Review* 41:963-80.
- Alwin, D. and R. M. Hauser
1975 "The decomposition of path effects." *American Sociological Review* 40:37-47.
- Boocock, S. S.
1972 *An Introduction to the Sociology of Learning*. New York: Houghton Mifflin.
- Bowles, S. and H. Gintis
1976 *Schooling in Capitalist America*. New York: Basic Books.
- Cicourel, A. V. and J. I. Kitsuse
1963 *The Educational Decision-Makers*. New York: Bobbs-Merrill.
- Clark, B.
1960 "The 'cooling out' function in higher education." *American Journal of Sociology* 65:569-76.
- Coleman, J. S., E. Q. Campbell, C. J. Hobson, J. McPartland, A. M. Mood, F. D. Weinfeld, and R. L. York
1966 *Equality of Educational Opportunity*. Washington, D.C.: U.S. Government Printing Office.
- College Entrance Examination Board
1968 *Effects of Coaching on Scholastic Aptitude Test Scores*. New York: College Entrance Examination Board.
- Cook, M. A. and K. L. Alexander
1977 "On the motivational relevance of educational plans." Unpublished paper, Department of Social Relations, Johns Hopkins University.
- DeBord, L., L. J. Griffin, and M. Clark
1977 "Race, sex, and schooling: insights from the Wisconsin Model of the early achievement process." *Sociology of Education* 50:85-102.
- Educational Testing Service
1957a *Cooperative School and College Ability Tests: Technical Report*. Princeton: Cooperative Test Division, Educational Testing Service.
1957b *Cooperative Sequential Tests of Educational Progress: Technical Report—Reading, Writing, Listening, Social Studies, Science, Mathematics*. Princeton: Cooperative Test Division, Educational Testing Service.
- Gintis, H.
1971 "Education, technology, and the characteristics of worker productivity." *American Economic Review* 61:266-79.
- Hauser, R. M., W. H. Sewell, and D. Alwin
1976 "High school effects on achievement." Pp. 309-41 in W. Sewell, R. M. Hauser, and D. L. Featherman (eds.), *Schooling and Achievement in American Society*. New York: Academic Press.
- Heise, D.
1972 "Employing nominal variables, induced variables and block variables in path analysis." *Sociological Methods and Research* 1:147-73.
- Heyns, B.
1974 "Social selection and stratification in schools." *American Journal of Sociology* 79:1434-51.
- Hut, M. and R. W. Morgan
1975 "Race and sex variations in the causes of the expected attainments of high school seniors." *American Journal of Sociology* 81:364-94.
- Karabel, J.
1972 "Community colleges and social stratification." *Harvard Educational Review* 42:521-62.
- Kerckhoff, A. C.
1976 "The status attainment process: socialization or allocation?" *Social Forces* 55:368-81.
- Kerckhoff, A. C. and R. T. Campbell
1977 "Black-white differences in the educational attainment process." *Sociology of Education* 50:15-27.
- McDill, E. L. and L. C. Rigsby
1973 *Structure and Process in Secondary Schools*. Baltimore: Johns Hopkins University Press.
- Persons, T.
1959 "The school class as a social system." *Harvard Educational Review* 29:297-318.
- Porter, J. N.
1974 "Race, socialization, and mobility in educational and early occupational attainment." *American Sociological Review* 39:303-16.
- Portes, A. and K. L. Wilson
1976 "Black-white differences in educational attainment." *American Sociological Review* 41:414-31.
- Ramsøy, N. R.
1965 "College recruitment and high school curricula." *Sociology of Education* 38:297-309.
- Rosenbaum, J. S.
1975 "The stratification of socialization processes." *American Sociological Review* 40:48-54.
- Sewell, W. H., A. O. Haller, and G. W. Ohlendorf
1970 "The educational and early occupational attainment process: replications and revisions." *American Sociological Review* 35:1014-27.
- Sewell, W. H., A. O. Haller, and A. Portes
1969 "The educational and early occupational attainment process." *American Sociological Review* 34:82-92.
- Shaycoft, M. F.
1967 *Project Talent, The High School Years: Growth in Cognitive Skills*. Pittsburgh: University of Pittsburgh Press.
- Thomas, G. E., K. L. Alexander, and B. K. Eckland
1977 "Access to higher education." Unpublished manuscript, Department of Social Relations, Johns Hopkins University.

MARX'S THEORY OF THE FALLING RATE OF PROFIT: TOWARDS A DIALECTICAL ANALYSIS OF STRUCTURAL SOCIAL CHANGE*

RICHARD P. APPELBAUM

University of California, Santa Barbara

American Sociological Review 1978, Vol. 43 (February):67-80

Conventional sociological theory is unable to account for endogenous change at the structural level. While various sociologists have attempted to develop a framework that would account for such change by abstracting formalized elements out of Marxist theory, these efforts—precisely because of their highly formal nature—are unsuccessful. Marx's theory itself seeks to explain change as built into the contradictory survival requirements of class societies conceived as closed social systems. The forces militating for change play themselves out with quasi-automatic necessity. The direction of change depends partly on the structural parameters and partly on the consciousness of individuals organized into social classes. It is argued that the utility of Marx's approach is tied both to his overall theoretical framework, and its rootedness in specific historical, social and economic conditions. This argument is illustrated with reference to Marx's theory of the falling rate of profit under conditions of competitive capitalism.

The past decade has seen an enormous increase in Marxist scholarship available in the English language. In addition to the numerous primary works and collections, key writings hitherto unavailable in English now have been translated and published.¹ The writings of Lenin are becoming available also, as are those of such early Hegelian Marxists as Lukacs.² A sizeable number of journals concerned with Marxist scholarship have appeared also in the past ten years.³ It is likely that

at no time since the first quarter of this century has there been a comparable renaissance of interest in Marxism in the United States.

This activity largely has been ignored within the confines of the established academic disciplines. The sociologist concerned with developing proficiency within a Marxist framework must look outside his field for direction and guidance. Since 1970 the principal journals of the profession have carried only some thirty articles that in any way address theoretical or empirical issues arising from the Marxist paradigm.⁴ As a direct consequence of

* Successive drafts and various versions of this paper have drawn upon the helpful criticisms of Tony Giddens, Alvin Gouldner, Harvey Molotch, James O'Connor, Bertell Ollman, Howard Sherman, Paul Sweezy, Erik Olin Wright, Morris Zelditch, and several anonymous referees. Responsibility for the final interpretation and argument is, of course, entirely my own.

¹ Principal among these is Nicolaus's translation of *The Grundrisse*, Marx's economic notebooks during the period 1857-8, appropriately subtitled "Foundations of the Critique of Political Economy." This work contains the most detailed exposition available of Marx's methodology, and is itself the groundwork for the writings which were to appear a decade later in *Capital*.

² Lukacs's influential writings during the 1920s were collected and published in German in 1968; the English translation was published in 1971, under the title, *History and Class Consciousness*.

³ For example, *Telos*, a journal heavily oriented towards phenomenological Marxism, first appeared

in 1968; *Radical America* in the mid-1960s; *Socialist Revolution* in 1970; the *Review of Radical Political Economics*, journal of the Union for Radical Political Economics (URPE) in the late 1960s; *The Insurgent Sociologist* in 1971; *Working Papers for a New Society* in 1973; *Working Papers on the Capitalist State* in 1973; *Crime and Social Justice* in 1974; *Theory and Society* in 1974; *Contemporary Crises* in 1976.

⁴ Four journals were searched for articles appearing between January 1970 and June 1977 which addressed the Marxist paradigm. The journals included *The American Sociological Review*, *American Journal of Sociology*, *Social Forces*, and *Social Problems*. The articles encountered were either theoretical in nature or concerned with empirical issues emerging from a Marxist perspective. Not all empirically-oriented articles cited Marx, but all examined Marxist hypotheses.

this ignorance, Marx's theory of social change has been misunderstood in American social science. While the most common tendency has been to vulgarize Marx as an economic determinist (Van den Berghe, 1963: 699-700), he also is treated as a propounder of an ill-defined dialectic (Van den Berghe, 1963; Schneider, 1971).

Nonetheless, there have been several efforts in recent decades to incorporate Marx into the sociological mainstream. These efforts generally take the form of distilling from Marx's writing that which is compatible with a functionalist approach to the social world—in particular, his concern with social conflict and change. Marx's theory is abandoned readily, while certain features of his dialectical method are extracted and preserved as heuristic principles. I have characterized such efforts as *formalist*. Formalism, in the sense employed here, has two related meanings: (1) the belief that Marx's *method* can be separated from his *theoretical framework* and fruitfully applied to another one (functionalism) which differs in its key conceptual categories and their interrelations; (2) the belief that it is possible to generate highly abstract propositions about the social world (in this case, derived from Marxist theory) that are transsituational and even transhistorical. Both of these meanings imply the tacit assumption that a social theory consists of various elements that logically can be separated both from one another and their empirical content, and subsequently recombined and re-applied in other contexts. Both misunderstand Marxism, and, more importantly, sacrifice the promise contained in Marx's approach to the study of political economy: a theory of endogenous structural social change which is neither determinist nor voluntarist. To elaborate these points is the task of this paper. I shall consider first several recent sociological treatments of Marx to illustrate the pitfalls resulting from the formalism inherent in their approach. I then shall examine one aspect of Marx's crisis theory in an effort to demonstrate that the utility of his method is bound up with his larger theoretical framework.

FORMALIST APPROACHES TO MARXIST ANALYSIS

Dialectics as Heuristics

Rytina and Loomis (1970:309) have claimed that "no American sociologist has tried to show that the dialectic, in a form that Marx would have recognized, is a useful approach to social science." There have, however, been some attempts to rescue the dialectic from Marxism and translate it to a more conventional sociological framework. Such efforts illustrate the pitfalls of formalism in the first sense of the preceding definition. We shall examine the efforts of Van den Berghe (1963) and Schneider (1971) in this regard.

Van den Berghe (1963:699) derives a "minimum dialectic approach" by eliminating what he perceives as unacceptable in Marx's method.⁵ This includes the latter's alleged "economic determinism," along with "his dualistic view of social reality," which "confused an empirical tendency for contradictions and conflicts to polarize into pairs of opposites, with a logical necessity to do so." What remains, according to Van den Berghe (1963:699), are two elements which appear both useful and valid:

- (1) Change is not only ubiquitous, but an important share of it is generated within the system; i.e., the social structure must be looked at, not only as the static framework of society, but also as the source of a crucial type of change.
- (2) Change of intra-systemic or endogenous origin often arises from contradiction and conflict between two or more opposing factors. These "factors" can be values, ideologies, roles, institutions, or groups.

This approach then is held to apply to three levels. One corresponds to Hegel's focus on ideas: "abstract but explicitly

⁵ Van den Berghe (1963:699) does not think much of "Hegelian-Marxist dialectic," since, after attempting to see "what can be usefully salvaged of the dialectic," he concludes, "admittedly not very much," although "the residual core is of great importance." Nonetheless, he does not hesitate to speak of the "dialectic method," a term he apparently uses interchangeably with "dialectic outlook."

formulated cultural concepts, viewed in isolation from concrete participants" (700). A second level corresponds to Marx's concern with social institutions: "the internal contradictions growing out of institutionalized processes of interaction" (700). A third level corresponds to a more general interest-group theory which views conflict pluralistically rather than dualistically.

Van den Berghe's "minimum dialectic approach" enables him to "transcend ad hoc eclecticism and . . . reach a balanced theoretical synthesis" (705) between Marxism and functionalism. He sees four commonalities between the two: (1) both are holistic in approach; (2) both involve conflict as well as integration—since conflict and integration imply one another (e.g., too much consensus or integration may impede social-system adaptation to changing circumstances, and hence result in disintegration and conflict); (3) both share a unilinear notion of social change in which future stages are contained within the present; and (4) both are fundamentally grounded in a dynamic equilibrium model of social organization. For dialectics—the synthesis, according to Van den Berghe—is the reestablishment of equilibrium after disturbance.

Schneider's reformulation resembles that of Van den Berghe in that he also extracts a dialectical perspective or bent expressed as a set of heuristic principles for an analytic framework capable of comprehending social conflict.⁶ The dialectical bent is legitimized by Schneider through demonstrating its presence in the classics (e.g., the nineteenth century Scottish historian, John Miller; Herbert Spencer; Pareto; Weber; Marx; Wundt; Veblen). This "bent . . . involved in some of the most fundamental insights in sociology" (1971: 669) reduces to three basic elements: (1) unanticipated/unintended consequences of purposive social action (Merton's latent

functions?), as a result of which man "comes to confront a world he may think he never made but which he did make," a notion therefore akin to the "phenomena of reification and alienation" (670); (2) heterogony of ends, whereby originally intended ends are displaced by means of secondary ones or are enlarged beyond original expectations (this resembles Hegel's "cunning of reason," where petty individual aims unintentionally contribute to history's larger purpose); and (3) paradox, or the "irony of sudden reversals," as for example when success leads to failure because initial adaptations turn out to be maladaptive in the long run.

Schneider then augments this historical review by distinguishing seven meaning-clusters derived from an examination of the sociological literature, which turn out to be largely reducible to the three basic elements distinguished above.⁷ He concludes with an assessment of the value of the dialectic as he has developed it, and expresses the hope that as a consequence of his study, we now have a "shrewd taxonomy not quite achieved previously and enhanced awareness of a certain kind of subtlety that attaches to particular kinds of social change . . ." (676).

Both Schneider and Van den Berghe have attempted to view formalistically the dialectic as a set of heuristic principles which might serve as a guide to more concrete scientific studies informed by these principles. Let us briefly summarize these heuristic principles in schematic form:

(I) *Holism*: the relevant unit of analysis is the social system, viewed as an interrelated configuration of differentiated elements. Van den Berghe (1963:701-2), however, mistakenly views Hegelian-

⁶ Schneider (1971:667), following Kaufman (1965), goes so far as to argue that "there is no dialectical 'method' to expound." Hence his use of the more restrictive terms "perspective" and "bent."

⁷ The meaning-clusters include (1) discrepancy between aims and outcomes (involving elements of unanticipated consequences and paradox); (2) goal shifts and displacements (heterogony of ends); (3) systems adaptations that ultimately prove dysfunctional (unanticipated consequences, paradox); (4) contradiction (which Schneider apparently equates with paradox); (5) the "contradictory logic of passion," which produces unanticipated consequences; (6) "development through conflict . . . stress," and (7) the dissolution of conflict "in a kind of coalescence of opposites," e.g., love and hatred (675-6).

Marxist analysis as tending to emphasize single-factor unidirectional causation, in contradistinction to functionalism, which he sees as arguing for "multiple and reciprocal causation."

(II) *Conflict*: this characterized, at least as much as integration, the relations among the parts of the relevant system. Conflict is a ubiquitous feature of social systems.

A. Social change results from conflict; hence, change is also a ubiquitous feature of all social systems.

B. Conflict is generated endogenously, i.e., as a result of conflict between two or more opposing forces within the social system (analogous to the Marxian notion of international contradiction).

C. Conflict often results as a latent function of social structures within the social system. This is because human actions and the consequent social institutions often eventuate in both unanticipated and/or unintended consequences and dysfunctions (from the long-run point of view of system adaptability).

(III) *Dynamic equilibrium*: this describes the movement of conflict-ridden, changing social systems through time. As conflicts are worked out, the system equilibrates at another level, until newly generated conflicts produce new disequilibria and new adjustments.

A. This dynamic equilibrium might also be described as a moving synthesis.

B. It implies an evolutionary view of social change in which future forms of social systems are contained within and reflect the working out of conflicts within the present forms.

Reduced to these principles, the dialectic becomes an empty truism. Common sense knowledge, dialectical or otherwise, tells us that social institutions are best conceived systematically; that conflict and change characterize many if not most social relations; that such conflict and change often occur because human actions produce unforeseen and undesired results; and that institutions, including societies, often evolve by fits-and-starts, rather than remaining forever stable. A dialectical approach must "overcome the predominantly arbitrary nature of unex-

plained historical events by deriving them from elements of their social structures" (Dahrendorf, 1964:100). Simply stated, it must account for the possibility of endogenous social change.

To address this issue, it is first necessary to distinguish endogenous from exogenous sources of change. In classical functionalist theory, the sources of social change were regarded ultimately as *external* to the social systems. Change was conceptualized as an internal response to changing external conditions. Thus, for example, Durkheim (1964:256-328) thought the increasing division of labor and the corresponding changes in the form of solidarity resulted from growing population ("moral density"). The growing population simultaneously increased pressure on scarce resources, thereby necessitating more efficient modes of social organization, and destroyed the social bases for widely shared collective beliefs. Left alone, isolated social systems, tending toward equilibrium or homeostasis (Appelbaum, 1970:65-80), were presumably highly stable over time. But mounting population (a biological, nonsocial condition) provides the cause of new system adjustments. There are other causes, but these are also external; e.g., contact with other cultures through warfare or trade. Classical functionalism, in regarding social systems as relatively stable, well-integrated configurations of elements in which each element contributes to the overall functioning of the whole, cannot account for internal sources of system change. Marxism, however, claims to be able to do just that. It is precisely this feature of dialectics that Schneider and Van den Berghe wish to preserve in their reformulations. What, in their views, are the possible sources of endogenous social change?

One set of possibilities, having little to do with the properties of social systems themselves, refer to the perspectives of the actors. These are changes that *appear* accidental or infelicitous to the actors involved. Such changes constitute several of Schneider's (1971) dialectical meaning-clusters (discrepancy between aim and outcome, contradiction, goal shifts

and displacements),⁸ as well as Van den Berghe's notion of latent functions. According to these formulations, our best-intended acts oftentimes betray us in ways unforeseen and unintended. Goals are not achieved because of unanticipated circumstances; we lose sight of our larger purposes along the way; we find that our actions, producing quite the contrary results from those originally intended, negate our original purposes, and cast us into ironic paradoxes. Schneider (1971: 675) treats "contradiction, oppositeness or opposition, paradox ('seeming contradiction'), negation, dilemma" as equivalent. In fact, he mistakenly identifies such inadvertent reversals with Hegel's "cunning of reason" (669, 671n, 673, 674) and Marx's "reification" (670). Such reversals are accidental, however, only from the perspective of the individual, who is presumed to act in a rational, goal-oriented fashion, yet who, for various reasons, fails to achieve that which was originally intended. Why the failure? This pivotal question is not addressed, but we may speculate. One possible explanation is lack of information. The individual actor, because of the complexities of social situations, simply does not possess all the knowledge s/he requires to fully anticipate all possible outcomes and thereby govern action appropriately. Another possibility has to do with changing circumstances. Actions appropriate to achieve goals under given conditions may produce quite different results when conditions change, if action is not suitably altered to reflect changing conditions. Whatever the reasons, the important feature of such accidental reversals, from the perspective of the actor, is their ironic or paradoxical nature.

The ironic or paradoxical quality of such seemingly accidental reversals is not, however, significant in terms of the task at hand: to derive a perspective which enables the *sociologist* to comprehend

change as necessarily internal to social structures. Whatever the sources of confusion for the actors involved, the sociologist is attempting to show that under given empirical conditions change *must* occur, and in a particular direction. Again, why should change be *necessary*—an embedded feature of social structures?

When Schneider (1971:673) talks of systems adaptations that prove dysfunctional, or when Van den Berghe (1963: 699) speaks of development through conflict (see also Schneider 1971:675), both men are referring to changes that result from occurrences within the social system which are subject to lawful explanation. While it is obvious that changes often occur through dysfunction and conflict, it is not clear from their examples why such change should be necessary. As an example of changes arising from dysfunction, Schneider (1971:674) offers highway construction which, although intended to relieve congestion, actually generates more traffic. As an example of the changes arising from conflict, we have interest group conflict of all forms (Van den Berghe, 1963:700). Is the former merely an instance of bad planning? The latter, the results of human rapaciousness? Why should highway planners fail to anticipate the larger consequences of their actions? Why should one group of people have interests antithetical to those of another? Are dysfunction and conflict *necessary* features of social systems?

Schneider (1971:673-4) offers a partial answer to this question when he speaks of "structures or forms that constitute relatively effective adaptations but stand in the way of more effective adaptations because of 'investments' already made." As an example, he offers Marx's theory of surplus value. After apologizing for using an instance he apparently regards as possessing doubtful empirical value, he nonetheless finds something suggestive in the form of Marx's reasoning:

A nuance arises as one thinks of a "system" context in which some specific product that emerges from and marks "success" turns about, so to speak, and leads to failure or system breakdown. . . . Surplus value does

⁸ Schneider's "heterogony of ends" would appear to be a specific instance of unanticipated/unintended consequences in which the pursuit of certain goals unexpectedly results in enlargement of those goals, their displacement by other goals, etc.

in a sense emerge from the success of capitalism. It unequivocally marks or signalizes that success. And yet in time, for economic (and class-psychological) reasons, this very product of surplus value brings about the downfall of capitalism. (673-4)

Why is this an example of short-run systems adaptations that prove dysfunctional in the long-run? Perhaps it is the narrow-mindedness of individual capitalists, who fail to appreciate that by garnering surplus value they will engineer their own defeat for economic and class-psychological reasons. In that case, we return again to the problem of inadequate information, and history can be cheated by a process of education. Or perhaps it is because an economic system based on surplus value—which appeared to constitute a relatively effective adaptation at one time (to what? for whom?)—proved to have deleterious effects in the long-run. Too many “investments” in capital and managerial forms (and perhaps class privilege), however, now impede individual capitalists from shifting to something new. Here the notion of institutional inertia must be invoked to account for the failure to respond effectively to changing conditions. Change is again conceived as resulting from an external source, in this case the ossification of social forms. In neither case, however, is a structural theory of social change offered. The first possibility ultimately locates change in individual perceptions, while the second invokes invariant and external constraints on social institutions. The example from Marx, however, does provide a clue, although in a somewhat different direction than that sought by Schneider. For Marx, surplus value does not merely emerge from, mark, nor signalize the success of capitalism; it rather is *required* by the internal logic of capitalist economic production. We shall pursue this central point in the final section of this paper.

Marxism as Conflict Theory

Both Schneider and Van den Berghe offer heuristic principles for a dialectical methodology. The formalism of this approach has been criticized insofar as it removes the promise Marxism purports to

offer as a theory of structural change. The same is true for those theorists who would restate Marx's work as a corpus of universal (i.e., transhistorical and transsituational) propositions concerning social conflict, the most prominent examples of which are contained in the writings of Dahrendorf (1959) and Turner (1973; 1974; 1975a; 1975b). Both writers are concerned to develop a general theory of social conflict in propositional form. Dahrendorf reformulates Marx by substituting authority for property as the defining characteristic of social class, and derives a theory of interest-group conflict. Insofar as his attention thereby shifts from the overall society to the level of the social group, his efforts are less central to the present discussion than are those of Turner (1975b:626), who limits himself to a propositional restatement of Marxism in order to better extract from him “what is theoretically useful and move on with the job of theory-building.” Turner (1975b:626) strongly argues that such theories “are most useful when stated at their most abstract level, for it is in this form that the debt of contemporary theorizing to these two German scholars [Simmel and Marx] becomes most evident.” Turner's approach, then, is self-consciously formalist in the second sense of that term as I have used it. His intent is to develop propositions of universal validity which can then be combined with other propositions from other sources, until a sociology of conflict is developed—an inventory of propositions that explain the source and evolution of social conflict in all forms. Turner's (1975b:621) propositions, derived from Marx, include the following:

- I. The more unequal the distribution of scarce resources in a system, the greater will be the conflict of interests between dominant and subordinate segments in a system.
- II. The more subordinate segments become aware of their true collective interests, the more likely they are to question the legitimacy of the inequality of distribution of scarce resources. . . .
- C. The more members of subordinate segments can communicate their grievances to each other, the more likely are they to become aware of their true collective interests.
- (1) The more ecological concentration of

members of subordinate groups, the more likely are they to communicate their grievances.

(2) The more subordinates have access to educational media, the more diverse the means of their communication, and the more likely they are to communicate their grievances.

What is striking about these propositions is their generality. It is not that they violate the spirit of their source, nor that there is no utility in rigorously stating one's hypotheses and conclusions. Rather, it is that these propositions, like the dialectical principles presented by Van den Berghe and Schneider, are so unspecified as to be largely empty. Consider (I). What "scarce resources?" Which "system?" Does or should this proposition apply with equal force to family groups, the military, all societies, religious groups, political parties? Marx's original formulations are concerned specifically with the relations of social classes in a class-based society. Why should a proposition based on this situation acquire universal validity? Marx himself invokes no abstract conflict of interests, which somehow is held to vary directly with the degree of maldistribution. Rather, he examines specific class conflicts, such as the historical struggle in England over the length of the working day, to spell out their relationship to such structural variables as the average rate of profit, the degree of monopolization, or the movement of the business cycle. Consider (IIC). What is meant by "ecological concentration?" This proposition is a generalized restatement of Marx's (Marx and Engels, 1848:345) contention in the *Communist Manifesto* that "the advance of industry, whose involuntary promoter is the bourgeoisie, replaces the isolation of the laborers, due to competition, by their revolutionary combination, due to association." But in the same document, Marx also notes that "this organization of the proletarians into a class, and consequently into a political party, is continually being upset again by the competition between the workers themselves" (343). Marx, in fact, states a number of conditions which mediate the level of class struggle in the *Manifesto*—a document intended as a

polemic, and which therefore tends to emphasize the inevitability of revolution. While Marx personally may have never doubted this long-run result, in his more systematic writings such as *Capital* he focused on the complexity of interrelated conditions that can modify the class struggle, including the degree of class consciousness, under specific conditions. Again, Turner is seeking a generalized restatement of that which Marx specifies as empirically existing circumstances. There is clearly truth in Turner's proposition; but one can immediately think of numerous exceptions, situations which mitigate or modify the law as stated. And it is precisely such mitigating circumstances that are of interest to Marx's theory, a theory of the historically concrete.

Marx himself, as is well-known, often refers to general laws. In Vol. 1 of *Capital*, he discusses the tendency of capitalist economic production to develop a reserve army of surplus labor in proportion to the expansion of capital. "This," says Marx (1867:644; emphasis removed), "is the absolute general law of capitalist accumulation." But Marx, in the very next sentence, qualifies his conclusion in a significant way. He observes that "like all other laws it is modified in its working by many circumstances." These circumstances are peculiar to the capitalist economic system under consideration. All circumstances are not equally important. Only within the framework of a suitable economic theory (Marx here offers a theory of the business cycle) does one know which circumstances to consider. Furthermore—and this point is central in distinguishing Marx's approach from those we have been considering—the theory itself is changed as the circumstances are altered through theoretically informed political practice.

While Marxist theory can fruitfully be restated in propositional form (for a useful example, see Gottheil, 1966), such propositions would have the following characteristics: (1) they would be interrelated within a logico-deductive theoretical framework, and acquire their significance only within that framework; hence, they would resist abstraction to broader, extratheoretical contexts; (2) they would be

highly contingent statements about empirical events, mediated by specifiable historical circumstances; (3) they themselves would change as the circumstances changed—theory and practice inform and modify one another. Turner's formalized propositions deliberately possess none of these characteristics. Marxism is the concrete analysis of concrete conditions. In this it differs from conventional social science, which, after the fashion of the hard sciences, seeks after universally valid, highly formalized laws (e.g., $f = ma$). Because of this, Marxism has a peculiar status as a generalizing science. It attempts to straddle the methodological schism between the idiographic and nomothetic approaches to social phenomena. While this makes for considerable ambiguity concerning the actual nature of a uniquely Marxist method, it also contains its unique promise: a theory capable of comprehending change as internal to actual sociohistoric systems, a function of forces which operate at the same time with lawful necessity while permitting a significant role to human agency in bringing about historical outcomes. To paraphrase Marx, science both interprets and changes history and is thereby itself changed in the process.

In the remainder of this paper, then, I shall argue that Marx's work offers a framework for comprehending endogenous structural change. The framework entails propositions concerning necessary or lawful relationships. Yet these propositions resist formalization after the fashion of the writers previously reviewed. This is due both to the highly contingent or specified nature of Marx's propositions and to the role assigned political practice. For Marx, political practice has the potential of modifying the conditions which give rise to the socioeconomic laws themselves. The laws governing social change, unlike the laws of physics or chemistry, do not permit prediction. They rather constitute a framework wherein a theoretically-informed, and hence effective, political practice is possible. In Marx's economic equations, the parameters themselves are treated as variable. The values of the parameters reflect historical conditions, in particular the state of

the class struggle. As such, they cannot be predicted. Consequently, one cannot predict the occurrence or outcomes of specific events with the aid of Marxian economics. These arguments will be illustrated by reference to Marx's theory of the declining rate of profit in capitalist economic production—a theory which might well appear to entail highly formal statements concerning inevitable tendencies subject to invariant laws and hence predictable outcomes. Although some Marxists read Marx in precisely this fashion, they are, I believe, mistaken. Marx's economic laws are not of this type.⁹

THE DIALECTIC OF PRAXIS AND STRUCTURAL CONSTRAINTS: MARX'S ANALYSIS OF THE FALLING RATE OF PROFIT

Marx analyzed the value of commodity production in terms of three elements: constant capital (C), the value of the means of production used up during the production process (primarily the depreciated value of machines, buildings, and raw materials); variable capital (V), the value of the labor-power applied to the production process (primarily the wage bill); and surplus value (S), the value of unpaid labor appropriated by the capitalist during the production process (workers' labor time beyond that which is socially necessary to sustain the standard of living of the working class). Surplus value is the key to capitalist economic production. It is the source of all profits, including those which are reinvested in enhanced productive capacity (capital accumulation). During the period of competitive capitalism, individual capitalists were under continual economic pressure to increase the efficiency of production—to produce commodities at lower unit costs. While this could be achieved by economizing on either of the two principal component costs of production, constant or variable

⁹ Marx's theory of the falling rate of profit is only one aspect of his overall theory of the crises of capitalist production. The purpose of the present discussion is *not* to present a thoroughgoing exposition and critique of crises theory, but rather to focus on Marx's original treatment of one principal source of crisis in order to better elucidate his method.

capital, Marx (1867:265; 1849:186) believed that in the long run the key to lowering production costs lay with mechanization. This meant increasing C relative to V . Thus, driven by the economic imperative to undersell one's competitors in order to remain afloat, individual capitalists would be driven to substitute increasingly efficient machines for human labor. Throughout the economy, therefore, there is a long-run tendency for what Marx termed the "organic composition of capital" to rise, as denoted by the symbol Q , where $Q = C/(C + V)$.¹⁰ This tendency, in turn, made it possible for the remaining workers to produce ever-larger quantities of goods with ever-decreasing labor time. As a consequence, there is a parallel tendency for the rate of surplus value (S'), defined as the ratio of unpaid to paid labor time (S/V), to rise as well.

How do these tendencies affect the overall rate of profit in capitalist economic production? Marx defines the rate of profit as the ratio of surplus value to total capital advanced, or

$$P = S/(C+V) \quad (1)$$

from which it follows algebraically that the rate of profit can be decomposed into two terms comprised of the rate of surplus value and the organic composition of capital:

$$P = S' (1 - Q) \quad (2)$$

where $S' = S/V$ and $Q = C/(C + V)$.

Marx (1867:449; Marx and Engels, 1848:338) argued that a rising organic composition was the hallmark of capitalist production. It follows that as Q ap-

proaches 1, $(1 - Q)$ approaches 0, with a consequent depressing effect on the overall rate of profit. On the other hand, inasmuch as the reason for mechanization in the first place is to increase S , the downward pressure on profitability resulting from rising Q will be partially offset. To the extent that S' rises as Q rises, the value of P is indeterminate. Marx (1867:247) was well aware of these considerations, but argued on logical grounds that as the organic composition reaches a high level, additional increases in productivity (hence S') are inadequate as a strategy to maintain profitability (see also 1857:338-40 for a crude mathematical proof).¹¹ Since capitalism is production for profit, once the overall rate of profit (or at least that obtaining in key economic sectors) drops below some minimally acceptable level, production ceases. Factories close down and an economic crisis ensues. The profit-maximizing strategy of individual capitalists has resulted in a profitability crisis for the class of capitalists as a whole. This is, for Marx, a structural imperative of capitalist economic production.¹² Yet despite the com-

¹¹ See Wright (1975: 37n) and Yaffe (1973:202) for a mathematical demonstration that "as the organic composition of capital rises, the rate of profit becomes progressively less sensitive to changes in the rate of exploitation [i.e., surplus value]" (Wright, 1975: 16). Marx (1867: 232-40) also details a number of empirical influences which may for a time counteract the tendency of the organic composition to rise, but these are not judged sufficient in the long run to mitigate the overall process.

¹² The crisis itself is an integral part of the dynamic of capitalist production. While it temporarily restores profitability through lowering the organic composition (see Yaffe, 1973: 205-6 for an elaboration), it does so by altering the framework of production itself, through contributing to the centralization of capital. Economic crises thus abet the transition from competitive to monopoly capitalism (Marx, 1867: 250-1; see Wright, 1975, for an excellent discussion of these processes). Some Marxists (Yaffe, 1973; Cogoy, 1973) argue that the falling rate of profit is ultimately the only source of crisis that follows necessarily from the logic of capitalist production, although of course other factors such as inadequate aggregate demand may shape specific crises. Others (Sweezy, 1968; 1974; Hodgson, 1974) argue that there is neither theoretical justification within Marx's work nor empirical evidence to support such a law, and that therefore, in Hodgson's (1974:65) words, "we are led to abandon the theory of the falling rate of profit, and along with it all vulgar

¹⁰ Marx generally speaks of the proportion or ratio of $c:v$; the organic composition of capital is expressed by some writers as c/v (e.g., Mattick, 1969; Mandel, 1968). We shall follow Sweezy's (1968) usage. He defines the organic composition as the ratio of constant capital to total capital advanced. Recent reformulations of Marx's theory of the declining rate of profit have argued that definitions involving only the terms C and V are inadequate, since V is itself dependent in part on the rate of surplus value (exploitation) (see, e.g., Cogoy, 1973; and Wright, 1975). In terms of the arguments raised in this paper concerning the relationships between the rate of profit, the rate of surplus value, and the organic composition of capital, it makes no difference which representation we employ. Our methodological conclusions are unaffected as well.

elling nature of such an imperative, its actual working-out depends on concrete sociohistorical circumstances. The declining rate of profit is a *tendency* that manifests itself within and through class struggle, and not a *law* which operates automatically outside of human practice.

In distinguishing tendency from law in Marx's theory, I am attempting to call attention to what I believe to be the central feature of interest in his method. Marx sought to avoid both the determinism of a completely materialist science, and the voluntarism of both idealist philosophy and utopian formulations. He achieved this by conceptualizing the material conditions of action as embedded within interrelated social, economic, and political structures. At the same time he regards human action itself as capable of modifying the underlying structures and hence the conditions of future action. In the example of the declining rate of profit during the period of competitive capitalism, the principal structures of analytic interest are economic. These include the factory system, organized such that social labor produces for the profit of the owners of capital; a market economy, in which production is for individual profit rather than collective utility; and overall economic organization predicated on competition among workers for jobs and capitalists for markets rather than on coordination and central social planning. These structures are organized neither as a congeries of random or accidentally related elements, nor as a determinate system in which all parts possess causal relationships with all others. Rather, the structures are conceived *dialectically*—a term we are now in a position to better understand.

The dialectic, as utilized by Marx in his economic analyses, refers to a mode of understanding empirical socioeconomic phenomena as *necessarily yet contradictorily* interrelated. Such interrelationships are unstable and hence must change, although within delimited bounds. The relationship between machines and workers,

conceived by Marx as a secular tendency for the organic composition of capital to rise, is a necessary yet contradictory one. It is *necessary* only under the historical "givens" of capitalist economic production: that production is for the profit of the capitalist-owner of the means of production; that capitalists compete with one another to sell their goods on the market; that workers are free to sell their labor to capitalists in exchange for a wage, rather than control over the productive process itself; that labor-power belongs to the capitalist for the duration of the working day, rather than merely that portion of the day necessary to sustain the worker. It is *contradictory* in that it is inherently unstable: it is impossible to sustain a rising organic composition of capital indefinitely without undermining the necessary profitability on which capitalist production rests. The various requirements of profitable production, dictated by the need for economic survival under capitalist economic organization, are mutually incompatible. Each individual capitalist must sell his commodities at the market price or below, or suffer a decline in sales to his competitors. To do so, he must continually seek ways to produce a larger volume of goods at lower unit costs, for that is what his competitors are doing. This requirement, in turn, engenders yet another; the need for productivity increases. Thus, the individual capitalist *must* economize, and increase the output per worker through the substitution of labor-saving technologies. Over the economy as a whole, however, this has the long-run consequence of lowering profitability and hence undermining production itself. Periodic crises are thus structured into capitalist production.

While the overall tendency of the contradiction can be deduced analytically from the "givens" of capitalist economic production, its concrete movement cannot. That depends on historically-specific circumstances. How is it that the basic conditions of the contradiction, the "givens" of capitalist economic production, can be modified? How is it possible to have a scientific theory of economic crisis that admits of "historically-specific circumstances?" These two questions are re-

notions of capital and capital accumulation." It is not the intention of the present paper to join that debate, although I find the arguments in favor of the necessity of the law to be unconvincing.

lated, for they both go to the root of the dialectic as a science of the historically concrete—a science that is not predicated on universal laws and predictive statements. We have thus far considered the dialectic as if it applied only to an external world of workers and owners, human and mechanical labor—in a word, the productive process. In other words, we have thus far treated the dialectic naturalistically. Marx, however, treated capitalist economic production, conceived as the relationship between constant and variable capital, as the unfolding of capitalist social relations among people rather than as naturalistic relations among land, working bodies, and machines. That is why his key conceptual building blocks (C, V, and S) cannot be regarded as independent variables in a predictive equation. Rather, they are *at once analytic categories for understanding structural relationships and tendencies, and political indicators of the degree of the class struggle*. This latter point is the key to understanding Marx's method, and can be illustrated by returning to the example of the declining rate of profit.

Equations (1) and (2) tell us a number of things about the rate of profit. First, they show that the rate of profit consists of specific relationships between constant capital, variable capital, and surplus value. Given values for C, V, and S, the rate of profit is directly determined. Thus, on this first level of understanding, the equations serve the customary role of such equations in all science: a formal representation of an empirical event. Like all such representations, they have the property of abstracting away from concrete phenomena to mathematical symbolism. The symbolic rendition, manipulated according to mathematical rules, permits deductions to be drawn.

This leads us directly to a second level of understanding. The postulated relationships among the terms in the equations constitute a theory about the world. In particular, in defining the rate of profit in this way, Marx is arguing that the rate of profit is directly proportional to the mass of surplus extracted from labor, and inversely proportional to the capital advanced. Inasmuch as Marx conceptualizes

surplus as unpaid labor time, this definition has far-reaching implications. It suggests that to the extent that living labor (V) is replaced by past or dead labor in the form of machinery (C), surplus value and hence profits will dry up, at least to the extent that increases in the rate of surplus value (S') prove inadequate to offset the necessarily rising organic composition (Q). Thus, given Marx's theory about the world, it is possible to deduce certain necessary relationships among the elements that constitute the world.

At these first two levels there is no difference between Marx's method and the methods of science in general—both the natural sciences and the conventional social sciences which model themselves on hard scientific methodology. The difference between Marx's method and scientific ones emerges when one asks the question, Under what conditions is a rising organic composition likely to actually outstrip a rising rate of surplus value? Such conditions are not deducible mathematically from the equations for the rate of profit; nor can they be derived from any other equations in Marx's theory. Rather, such conditions are the result of human activity; in particular, the state of the struggle between labor and capital. The locus of this struggle is denoted by S, V, and C.

Surplus value (S), for example, is the arena of the struggle between workers and capital over the duration of the working day and the intensity of the labor process. That is why Marx (1867: chaps. 7–9), after introducing the concepts of C, V, and S in *Capital*, immediately proceeds to a discussion of the concept of absolute surplus value and a lengthy historical exposition of the struggle over the working day (1867: chap. 10). That is also why he follows the introduction of the concept of relative surplus value (1867: chap. 12) with documented discussion of the intensity of the labor process in modern industry (1867: chaps. 15, 17). The outcome of the class struggle over the disposition of the surplus is not derivative from the theoretical conceptualization of capitalist production, although it is certainly shaped by the conditions of production. For example, the concentration of workers in factories

facilitates their "revolutionary combination, due to association" (Marx and Engels, 1848:345). Rather, the outcome of that struggle, which occurs *within* the framework of structural conditions represented in formal equations, is ultimately a product of unique historical circumstances. The degree of working-class consciousness and political mobilization; the level of theoretical understanding of the working class and its leaders; the ability of capital to extract surplus from foreign workers to the advantage of domestic workers in relatively industrialized nations; the extent to which monopolization alters the ability of major capitalists to amass large surpluses while retarding the increase of the organic composition of capital—these and other factors will affect the disposition of the surplus according to empirical conditions.

Variable capital (V) draws our attention to another set of struggles: those having to do with the wage bill and workers' subsistence in general. The numerical values taken on by this term will reflect the degree to which labor is able to resist capital's efforts to depress its value. These efforts, as we have seen, take the form of cheapening the means of subsistence, reducing all labor to unskilled detail labor (and hence reducing the costs of subsistence by making it possible for all members of the family to work), depressing wages below their value, and shifting production to colonies with a ready source of cheap labor. Labor's success in this struggle depends, in turn, on its degree of organization and militancy. It depends upon the strength of unions, the access of the labor organizations to parliamentary institutions, and the degree of internationalization of the working-class movement (which means, for Marx, the Communist party). It also depends on the degree of tolerance for organized labor on the part of capital: whether or not labor is ruthlessly suppressed, legitimately accepted as part of the political structure, or even co-opted into its leadership. It is precisely because none of these conditions is derivative directly from the economic conditions that Marx stressed the central importance of organizing the international Communist movement as the only viable

means to shape those conditions to the advantage of the working class, and devoted much of his political life to doing so (see, for example, Marx and Engels, 1848: 345–6, 361–2).

The final set of struggles, those having to do with constant capital (C), concern the extensiveness of capitalist economic relations. As we have seen, the numerical values taken on by this term reflect in part technological conditions (e.g., the extent to which labor-saving technologies are capital-saving as well), in part the organization of production (e.g., the rate of capital turnover), and in part the ability of capitalists to extend capitalist economic relations abroad (e.g., foreign investment, which permits the importation of cheap raw materials and machinery produced with low-paid foreign labor). The value of constant capital in Marx's equations will also reflect the sociopolitical organization of capitalists as a class. It will reflect the intensity of competition among capitals, the degree of cooperation and centralization among the owners of large capital, and the extensiveness of capitals' ideological and political hegemony. This, in turn, draws attention to the role of the state in securing capitalist economic relations—the degree to which it is relatively autonomous or serves simply as an instrument of a unified ruling class, its executive committee (Marx and Engels, 1848:337). Again, such circumstances cannot be directly predicted on the basis of underlying economic conditions, although such conditions limit and shape the range of possible choices available to capitalists, acting individually or in concert as a class.

The struggle over the disposition of the surplus, workers' share in output, and capitalists' control over the political and ideological spheres and the productive process are merely different aspects of the class struggle. That struggle is not purely economic, although it depends to a large extent on economic conditions and affects those conditions most directly. The class struggle, as seen by Marx, also moves at the political and cultural levels. The possibilities of delegitimation of the state and dereification of both popular and scientific culture flow from economic struggles, and shape those struggles. That is

why theory itself becomes a material force when it has "seized the masses" (Marx, 1843:18).

There is nothing automatic about the processes of social change. This is true of the "stages of societal development" often attributed to Marx, as well as of more historically bounded economic "laws" (e.g., that of the falling rate of profit under capitalism). The movement of concrete societies occur within well-defined structural limits. Those limits are given for capitalist forms, within Marxist political economy, by the hypothesized relationships among the parameters C , V , and S . But those limits can be changed, the relationships among parameters themselves can change in value independently of their necessary connection within formal equations. This is because C , V , and S , while serving as economic parameters, are ultimately conceptualized by Marx as signifying social relations, of which the quantitative economic measures (hours of labor time, price) are merely surface indicators. Social relations can be altered, within bounds. Those bounds are, for Marx, first and foremost the structural conditions of economic production. The structural conditions generate problems (contradictions), and set limits to the solution of those problems. The solutions to the problems of competitive capitalism historically involved the emergence of monopoly capitalism and economic imperialism. These solutions, in turn, engender their own structural frameworks, replete with their own historically-specific contradictions demanding new solutions if the stability of the system was to be maintained (see Baran and Sweezy, 1966; Mattick, 1969; Wright, 1975; and O'Connor, 1973). The likelihood and efficacy of any economic solution depend, in large part, on the legitimacy accorded to the growing state intervention with its mounting economic costs and on political consciousness and class militancy in general. Crises of legitimation, dereification, class organization and struggle may grow out of adverse economic conditions or equally adverse solutions to such conditions, but they are not reducible to economic factors. The future cannot be predicted from within a

Marxist framework; it can only be shaped. It can be shaped scientifically to the extent that actions are based on adequate theoretical understanding of economic, political, and cultural structures. The terms "scientific" and "theoretical," however, must be understood in the limited sense developed in this paper. Marxist theory is a theory of structural constraints and probable tendencies which can themselves be shaped and altered by human praxis.

As a theory of social change, Marxism thus resists formalization after the fashion of the writers considered in the first part of this paper. It makes little sense to speak of a dialectical perspective on endogenous structural change, apart from the Marxist theoretical framework as applied to specified socioeconomic conditions. While it is certainly possible that a dialectic might be developed apart from such notions as surplus value, production for exchange, and class conflict, it is by no means self-evident that such a dialectic would remain fruitful for sociological analysis. The arguments and examples thus far offered to this end are unconvincing. A more promising approach would appear to lie within the Marxist paradigm itself. One might begin with the conceptual categories of Marxism and apply them empirically to contemporary conditions. As indicated previously, there is evidence that Marxist *theory* is gaining increased acceptance in the United States. The level of theoretical discussions is considerably advanced over that of only a decade ago. The level of empirical research, informed by that perspective, however, remains minimal. Should a comparable upsurge in such research occur, then we can expect useful revisions in theory itself. But theoretical and methodological advances are tied to theoretically informed research. They do not result from formalized borrowings, however more palatable such an approach may be to American social scientists.

REFERENCES

- Appelbaum, Richard
1970 *Theories of Social Change*, Chicago: Markham.

- Baran, Paul A. and Paul Sweezy
1966 *Monopoly Capital*. New York: Monthly Review.
- Cogoy, Mario
1973 "The fall of the rate of profit and the theory of accumulation: a reply to Paul Sweezy." *Bulletin of the Conference of Socialist Economists* 52-67.
- Dahrendorf, Ralf
1959 *Class and Class Conflict in Industrial Society*. Stanford: Stanford University Press.
1964 "Towards a theory of social conflict." Pp. 98-111 in Amitai and Eva Etzioni (eds.), *Social Change*. New York: Basic Books.
- Durkheim, Emile
1964 *The Division of Labor in Society*. New York: Basic Books.
- Gottheil, Fred M.
1966 *Marx's Economic Predictions*. Evanston: Northwestern University Press.
- Hodgson, Geoff
1974 "The theory of the falling rate of profit." *New Left Review* 84: 55-82.
- Kaufmann, Walter
1965 *Hegel*. New York: Doubleday.
- Lukacs, Georgy
1971 *History and Class Consciousness*. Cambridge, Ma.: Massachusetts Institute of Technology Press.
- Mandel, Ernest
1968 *Marxist Economic Theory*. 2 Vols. New York: Monthly Review.
- Marx, Karl
[1867] *Capital*, Vol. 1. New York: International Publishers.
[1843] "Critique of Hegel's philosophy of right." Pp. 11-23 in Robert C. Tucker (ed.), *The Marx-Engels Reader*. New York: Norton.
[1849] "Wage labour and capital." Pp. 167-90 in Robert C. Tucker (ed.), *The Marx-Engels Reader*. New York: Norton.
[1857] *Grundrisse*. New York: Vintage.
1973
- Marx, Karl and Frederick Engels
[1848] *Communist Manifesto*. Pp. 331-62 in 1972 Robert C. Tucker (ed.), *The Marx-Engels Reader*. New York: Norton.
- Mattick, Paul
1969 *Marx and Keynes: The Limits of the Mixed Economy*. Boston: Porter Sargent.
- O'Connor, James
1973 *The Fiscal Crisis of the State*. New York: St. Martin's.
- Rytina, Joan Huber and Charles P. Loomis
1970 "Marxist dialectic and pragmatism: power as knowledge." *American Sociological Review* 35: 308-18.
- Schneider, Louis
1971 "Dialectic in sociology." *American Sociological Review* 36: 667-78.
- Sweezy, Paul
1968 *The Theory of Capitalist Development*. New York: Monthly Review.
1974 "Some problems in the theory of capitalist accumulation." *Monthly Review* 26:1.
- Turner, Jonathan
1973 "From utopia to where? A strategy for reformulating the Dahrendorf conflict model." *Social Forces* 52: 236-44.
1974 *The Structure of Sociological Theory*. Homewood, Il.: Dorsey.
1975a "A strategy for reformulating the dialectical and functional theories of conflict." *Social Forces* 53: 433-44.
1975b "Marx and Simmel revisited: reexamining the foundations of social conflict theory." *Social Forces* 53: 618-27.
- Van den Berghe, Pierre
1963 "Dialectic and functionalism: toward a theoretical synthesis." *American Sociological Review* 28: 695-705.
- Wright, Eric Olin
1975 "Alternative perspectives in the Marxist theory of accumulation and crisis." *The Insurgent Sociologist* 6: 5-39.
- Yaffe, David S.
1973 "The Marxian theory of crisis, capital, and the state." *Economy and Society* 2: 186-232.

MANUSCRIPTS FOR THE ASA ROSE SOCIOLOGY SERIES

Manuscripts (100 to 300 typed pages) are solicited for publication in the ASA *Arnold and Caroline Rose Monograph Series*. The Series welcomes a variety of types of sociological work—qualitative or quantitative empirical studies, and theoretical or methodological treatises. An author should submit three copies of a manuscript for consideration to the Series Editor, Professor Robin M. Williams, Jr., Department of Sociology, Cornell University, Ithaca, New York 14853.

RESIDENTIAL SEGREGATION BY OCCUPATION AND RACE IN TEN URBANIZED AREAS, 1950-1970 *

ALBERT A. SIMKUS

University of Wisconsin, Madison

American Sociological Review 1978, Vol. 43 (February):81-93

Earlier studies by Duncan and Duncan (1955), Wilkins (1956), and Uyeki (1964) examined occupational residential segregation in a total of ten U.S. urbanized areas in 1950. The present study is a partial replication of these previous investigations, directed at measuring the changes in residential segregation in these same urbanized areas during the 1950s and 1960s. Changes in the relationship between racial segregation and occupational residential segregation are also examined. Occupational residential segregation between most occupational categories was found to have slightly increased during the 1950s. During the 1960s the degree of residential segregation between service workers and laborers vis-à-vis those in the higher occupational categories decreased, while segregation between persons in the higher categories remained much the same. Indexes of occupational residential dissimilarity calculated within and between racial groups reveal the degree to which gross occupational residential segregation was due to racial residential segregation and differentials between occupational categories in racial composition. These indexes also show that the degree of racial residential segregation depended somewhat upon the respective occupations of the whites and nonwhites whose residential distributions were compared. In 1960, the degree of racial residential segregation was slightly lower between whites and nonwhites in the lower categories. Between 1960 and 1970 nonwhites in the highest occupational categories became slightly less segregated from whites, while whites and nonwhites in the lowest occupational categories became slightly more segregated.

The degree to which members of different occupational categories are residentially segregated from each other is an important aspect of the long standing and general interest in the amount of segregation between members of different social strata. Such segregation is not only interesting in itself, it is often assumed that the degree and form of residential segregation may be a cause, correlate, or consequence of other aspects of social structure. In an often cited and reprinted article, Duncan and Duncan (1955a) demonstrated that the relative dissimilarities of

the residential distributions of members of eight major occupational categories conformed closely to a simplex structure such as might be expected from the differences in the general socioeconomic status of those occupations. The Duncans also pointed out the isomorphism between the pattern of dissimilarities between occupational categories in terms of residential distribution and the dissimilarities between occupational categories in terms of intergenerational occupational mobility. The same simplex structure of differences between occupational categories can also be observed in patterns of assortative marriage (Blau and Duncan, 1967).

The original study by Duncan and Duncan (1955a) was based on the occupational residential distribution over tracts and zone sectors in the tracted portions of Chicago and adjacent areas in 1950. Wilkins (1956) extended this line of analysis in a study of eight other urbanized areas in 1950 and found similar patterns of such segregation. The Duncans' findings were also replicated by Uyeki (1964; 1975) in studies of the metropolitan district of Cleveland in 1940 through 1970. Uyeki

* This is a revised version of a paper presented at the annual meeting of the American Sociological Association in New York, 1976. I wish to thank Kuanjeng Chen, William Frey, Karl Taeuber, Hallman Winsborough, and an anonymous reviewer for comments on an earlier draft. Of course, I am responsible for any errors which may remain. Barbara Aldrich and David Dickens provided assistance in programming. Computing facilities at the Center for Demography and Ecology were supported by the National Institute of Child Health and Human Development (NICHD) grant 1PO1-HD05876. This research was also supported by a National Institute of Child Health and Human Development training grant (1T32-HD07014).

found consistency between the patterns of residential segregation in the Chicago area and the Cleveland area in 1950. He also reported little change in the degree of segregation over zone sectors between 1950 and 1960.

The present study provides a partial replication of the three earlier studies of occupational residential segregation in a total of ten urbanized areas done by Duncan and Duncan (1955a), Wilkins (1956), and Uyeki (1964). Whereas these earlier studies focused largely upon the ordinal hierarchy implied by the relative degrees of dissimilarity between each pair of occupations, this study is primarily concerned with changes between 1950 and 1970 in the magnitude of segregation between members of each occupational category. Since the degree of occupational residential segregation is confounded by racial residential segregation and the different racial compositions of occupational categories, the analysis for 1960 and 1970 is carried out both for the total population and also within and across categories of race. In short, this study is directed at answering the following questions regarding residential segregation in these ten urbanized areas:

(1) What changes occurred in the degree of gross occupational residential segregation during the 1950s and 1960s?

(2) To what degree can the amount of gross occupational residential segregation in 1960 and 1970 be attributed to racial residential segregation and the varying racial compositions of occupational categories?

(3) What changes occurred during the 1960s in the level of occupational residential segregation within categories of race?

(4) What was the degree of residential segregation between persons who differed in both occupation and race, and how did this aspect of residential segregation change during the 1960s?

Data and Method of Analysis

The sources of data for the present study were the U.S. Census Population Summary Tapes for 1960 and 1970 (U.S. Bureau of the Census, 1972a; 1972b;

1972c).¹ The urbanized areas studied were Hartford, Syracuse, Chicago-Gary,² Cleveland,³ Columbus, Indianapolis, Fort Worth, Atlanta, Memphis, and Richmond. In each year considered, tracts were included in the analysis on the basis of whether or not a majority of the tract population resided within the appropriate urbanized area as defined in that year.⁴

¹ The tract data for 1970 was obtained from the Fourth Count, File A tapes for the appropriate states. Indexes calculated from the census tapes are necessarily somewhat higher than the population parameters for two reasons. The first reason is that the occupational information was obtained from a 20% sample. In an extreme case, if there were one resident of a certain type in each of five tracts, we would expect one of these five residents to be included in the sample. After correcting for the sampling fraction, the census would indicate five persons of that type residing in one tract and no such residents in the remaining four tracts, thus inflating the apparent unevenness of the distribution of these persons over the five tracts. The second reason is that information was suppressed for whites, blacks, or Spanish Americans whenever there were fewer than 15 respondents of that ethnic category (after allocations). Thus, if there were fewer than three blacks sampled from predominantly white tract, the data would show no blacks in that tract, and thus would increase the apparent degree of racial segregation.

² The area examined by Duncan and Duncan was the tracted area for Chicago-Gary and adjacent areas in 1950 (coextensive with the Chicago metropolitan district as delineated in 1940).

³ The area examined by Uyeki was the tracted area within the boundaries of the Cleveland metropolitan district, as delineated in 1940.

⁴ In this study, the degree of segregation is computed for each urbanized area on the basis of the urbanized area boundaries as defined at the time in question. The objection may be made that comparisons across time are being made on the basis of geographic areas which are not comparable, since both the boundaries of the urbanized areas and the boundaries of tracts within the urbanized areas changed from decade to decade. The subject of interest is the degree of residential segregation within urbanized areas—regardless of the expansion of the geographical boundaries within which that segregation took place. Limiting study to the boundaries of an "old" urbanized area would involve excluding not only the real urbanized area but also many of the people who participate in the social organization of that urbanized area. Tracts are drawn up so as to maintain roughly equivalent population sizes—not areal sizes. Since tracts are often split in order to maintain a roughly equal number of residents within each tract, changes in tract boundaries within the urbanized areas serve to maintain consistency in the segregation indexes calculated over units assumed to contain a consistent population size. In making comparisons across time, we must assume that tract

As in the earlier studies, the degree of occupational residential segregation was measured using the index of dissimilarity (d), computed on the basis of the residential distribution of employed males in each of eight major occupational categories, over tracts.⁵ The major occupational categories employed were: (1) professionals, (2) managers, officials, and proprietors, (3) salesmen, (4) clerical workers, (5) craftsmen, (6) operatives, (7) service workers (private household workers not included), and (8) nonfarm laborers.⁶ In the analysis involving categories of race, the population was dichotomized as white and nonwhite. The magnitude of the index of dissimilarity is dependent upon the size of the areal unit used in the analysis. Although it would be interesting to examine indexes calculated on the basis of smaller areas such as blocks, tracts are the small-

est units for which census tabulations of occupation by residential location are available. For this reason the present analysis and those studies being replicated deal with segregation across census tracts.

In order to display parsimoniously the results of the analysis, only the unweighted means of the indexes averaged over all ten urbanized areas will be presented. This procedure seems justified since the patterns and changes observed were similar in each of the urbanized areas examined. Of course, these ten urbanized areas do not constitute a representative sample of any larger universe of urbanized areas and the mean values reported should not be interpreted as estimates of the mean values for all U.S. urbanized areas. Differences between these urbanized areas are not examined in this paper. The pattern of segregation was similar across areas in 1960 and became more similar by 1970.

boundaries have not been changed so as to conform to occupational or racial residential boundaries to a greater or lesser degree.

⁵ The calculation of the index of dissimilarity is described in detail in Duncan and Duncan (1955a), and its merits as an index of segregation are discussed in detail in another paper by Duncan and Duncan (1955b) and by Taeuber and Taeuber (1965). As Cortese et al. (1976) have reminded us recently, another consideration regarding the use of the index of dissimilarity is that this index has a value of zero when members of two categories are distributed perfectly evenly over tracts. However, if the distribution of members of both categories is random and the average number of members of each group residing in each tract is quite small, the expected value of the index is greater than zero (see also Taeuber and Taeuber, 1965). In most instances the number of residents per tract (1,000–1,500) is large enough that the expected value of the index is close to zero. But in computing indexes between very small subgroups of the population, such as between black managers and black salespersons in an urbanized area with a small proportion of black residents, this property of the index should be taken into account. The suggestion by Cortese et al. (1976) that the index be standardized as they describe has not been followed. The Taeubers (1976) have argued that this is a dubious improvement upon the measure. In the present case such standardized indexes would be even more sensitive to differences in the racial and occupational compositions of the urbanized areas than are the unstandardized indexes.

⁶ Changes between 1950 and 1970 in the census occupational classification system (U.S. Bureau of the Census, 1962; 1972d) may have had some effect upon the sizes of the indexes of occupational residential dissimilarity in each decade; however, these effects should have been relatively small.

Findings: Changes in Gross Occupational Residential Segregation

In general, the degree of occupational residential segregation appears to have slightly increased in these urbanized areas during the 1950s. The indexes of dissimilarity (d) of the residential distributions of persons in each occupational category averaged over the ten urbanized areas in 1960 are displayed above the diagonal in Table 1. Below the diagonal are shown the changes in these indexes between 1950 and 1960 (1960–1950). Of the twenty-eight comparisons between the eight occupational categories, the degree of residential dissimilarity increased in twenty-three cases, decreased in three, and remained the same in two. The size of these changes was relatively small. The average index increased by 2.3 points, the largest increases were of 6 points, and the largest decrease was of only 2 points. The largest changes (small as they were) involved the distribution of clerical workers. The three decreases involved a lessening of the segregation between clerical workers and the three lowest occupational categories, while the largest increases involved segregation of clerical workers vis-à-vis managers and salesmen.

Table 1. Indexes of Residential Dissimilarity among Employed Males in Major Occupational Groups, Averaged across the Ten Urbanized Areas, 1960 and Change 1950-1960*

Major Occupational Group	Major Occupational Group							
	Prof.	Mgr.	Sales	Cler.	Craft	Oper.	Serv.	Labr.
Prof.		16	16	29	36	47	48	57
Mgr.	3		14	30	35	47	49	59
Sales	1	2		24	31	43	45	55
Cler.	3	6	6		17	25	31	41
Craft	0	3	2	1		20	33	40
Oper.	1	5	4	-1	3		23	26
Serv.	2	5	3	-1	2	0		23
Labr.	0	5	2	-2	3	1	2	

* Above diagonal, 1960; below diagonal, change from 1950 to 1960. The values for 1950 were taken from Duncan and Duncan (1955a; 1955b), Wilkins (1956), and Uyeki (1964).

During the 1960s there was a different pattern of change from that observed in the previous decade. The gross indexes for 1970 are shown in Table 2 above the diagonal, and the changes between 1960 and 1970 (1970-1960) are shown below the diagonal. The degrees of segregation between each of the six highest occupational categories remained virtually the same; they changed by an average of less than 1 point. However, segregation between the two lowest categories (service workers and laborers) and the six higher categories significantly decreased. The indexes of dissimilarity between service workers and the six higher categories decreased by an average of 7.6 points, while the indexes between laborers and the six highest categories decreased by an average of 11.0 points.

As a result of these changes, by 1970 service workers were less segregated from the five highest occupational categories than were operatives. Some such inversions in the rankings of service workers and operatives in terms of the degree to

which they were segregated from the higher occupational categories were observed in 1950 by the Duncans, Wilkins, and Uyeki, but by 1970 such inversions were observed consistently rather than as exceptional cases. These inversions intrigued the authors of the earlier studies. The Duncans (1955a) speculated that service workers were less segregated from the white-collar categories than were operatives because the service worker category includes apartment building janitors who may live at their place of work. However, Wilkins (1956) found no relationship between the presence or strength of these particular inversions and the proportion of service workers in each urbanized area who were janitors. To the Duncans, these inversions were particularly interesting because the same inversions were observed in terms of the dissimilarities between outflow percentages in tables of intergenerational occupational mobility. Thus, the ordering of occupational categories implied by residential segregation was more consistent with the ordering

Table 2. Indexes of Residential Dissimilarity among Employed Males in Major Occupational Groups, Averaged across the Ten Urbanized Areas, 1970 and Change 1960-1970*

Major Occupational Group	Major Occupational Group							
	Prof.	Mgr.	Sales	Cler.	Craft	Oper.	Serv.	Labr.
Prof.		16	16	28	36	45	40	46
Mgr.	0		14	30	36	46	42	48
Sales	0	0		27	34	44	39	45
Cler.	-1	0	3		20	25	22	29
Craft	0	1	3	3		20	25	29
Oper.	-2	-1	1	0	0		19	19
Serv.	-8	-7	-6	-9	-8	-4		19
Labr.	-11	-11	-10	-12	-11	-7	-4	

* Above diagonal, 1970; below diagonal, change from 1960 to 1970.

implied by occupational mobility than with the ordering of the occupations in terms of either education or income.

Between 1950 and 1970 the service worker category improved its position vis-à-vis the other categories in terms of educational attainment but not in terms of income. In 1950 the ratio of national median years of school completed for male service workers to that of male operatives was .99 (8.8 to 8.9 years). By 1960 this ratio was 1.01 (9.7 to 9.6), and by 1970 it was 1.02 (11.2 to 11.0). The ratio of median wage and salary income of male service workers to that of operatives was .82 in 1950, .76 in 1960, and .74 in 1970 (U.S. Bureau of the Census, 1952; 1961; 1973). In short, service workers were clearly less segregated from persons in the highest occupational categories than were operatives, even though operatives were closer to the higher categories in income. This finding reinforces the impression that such residential segregation is due to more than a simple function of differences in financial resources, and may be affected by status group differences in life style, aspirations, place of work, or intergroup affinity.

Residential Segregation by Occupation and Race

Calculating indexes of residential dissimilarity between persons categorized by both occupation and race, it is possible to assess (1) the degree of occupational residential segregation free of the confounding effects of racial segregation; (2) the effects of racial and occupational segregation acting in combination. A matrix displaying the dissimilarities in the residential distributions of employed males categorized by both occupation and race averaged over the ten urbanized areas is presented in Table 3. The triangular submatrix above the diagonal shows the values for 1960, and the values for 1970 are shown below the diagonal. Note that each of the triangular submatrices above and below the diagonal may be further subdivided into three parts: (1) a triangular matrix of indexes referring to occupational segregation among whites, (2) a square matrix of indexes referring to

segregation by both race and occupation, and (3) a triangular matrix of indexes showing occupational segregation among nonwhites.

Unfortunately, it is difficult to grasp an overall picture of the large number of social distances represented in each half of the matrix in Table 3. In the belief that one picture may be worth 120 numbers, it seemed useful to reduce this information to a somewhat simpler form. Smallest-space analysis (Lingoes, 1973) was used in order to obtain a graphic representation of the residential social space in these urbanized areas.⁷ Smallest-space analysis provides a means of reducing a matrix of distances (such as the indexes of residential similarity) between objects (such as the 16 categories of occupation by race) to an n -dimensional set of coordinates in which the Euclidian distances between the points reflect the original ordinal relationships among distances to the maximum possible degree. Although the derived coordinates are based upon ordered relations rather than metric distances, the very large number of ordered comparisons involved in a large matrix of distances results in a tightly constrained solution with a metric interpretation.

In the present case, the triangular matrices of indexes of residential dissimilarity in each half of Table 3 were used as input for the Guttman-Lingoes (Lingoes, 1973) program Misissa-1(M), resulting in two-dimensional representations of the residential social distances between the various pairs of categories of occupation by race.⁸ Similar previous uses of smallest-space analysis in conjunction with indexes of dissimilarity have involved social dis-

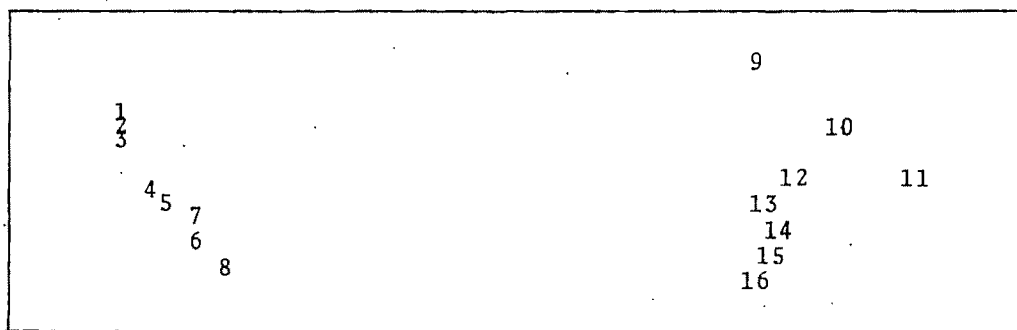
⁷ An excellent brief introduction to smallest-space analysis and its use in indicating social distance has been presented by McFarland and Brown (1973).

⁸ Solutions involving both more than two and less than two dimensions were also obtained; however, the two-dimensional solution was judged to be most satisfactory. A one-dimensional solution results in combining the racial and socioeconomic dimensions, while three or higher dimensional solutions produce dimensions which are difficult to interpret and which provide virtually no improvement in the fit between the original matrix and the derived distances. Kruskal's stress was less than .02, and the coefficient of alienation was less than .03 for each of the two-dimensional solutions in Figures 1 and 2.

Table 3. Indexes of Residential Dissimilarity Calculated between Employed Males Categorized by Occupation and Race, Averaged over the Ten Urbanized Areas, 1960 and 1970*

White										Nonwhite									
White					Nonwhite					White					Nonwhite				
Prof.	Mgr.	Sales	Cler.	Labr.	Prof.	Mgr.	Sales	Cler.	Labr.	Prof.	Mgr.	Sales	Cler.	Craft	Oper.	Serv.	Labr.	Serv.	Labr.
White																			
Prof.	15	16	28	36	44	39	47	85	90	93	89	88	89	88	89	89	88	89	88
Mgr.	15	14	28	35	44	39	46	86	89	94	89	87	89	87	89	88	88	88	88
Sales	15	14	22	31	39	34	42	85	89	93	88	86	87	86	87	87	87	87	87
Cler.	26	28	25	17	23	20	30	83	87	92	86	84	85	84	85	85	85	84	84
Craft	36	36	33	19	15	21	26	84	87	91	86	83	84	83	84	84	83	83	83
Oper.	43	43	40	25	14	20	20	83	85	90	84	81	81	81	81	82	80	80	80
Serv.	34	35	32	18	18	20	26	81	85	89	84	81	81	81	81	81	81	81	81
Labr.	39	41	38	24	20	19	21	82	84	88	82	79	79	79	79	80	78	78	78
Nonwhite																			
Prof.	77	80	79	77	79	76	77	43	45	45	37	43	46	45	46	45	49	49	49
Mgr.	84	86	85	84	85	84	84	43	46	43	34	38	37	38	37	38	41	41	41
Sales	89	90	89	89	88	87	87	45	46	46	36	38	38	37	38	37	40	40	40
Cler.	87	88	87	85	84	83	84	38	39	37	26	26	26	29	26	29	33	33	33
Craft	87	87	86	83	82	81	81	46	42	40	26	17	17	21	17	21	24	24	24
Oper.	88	89	88	86	84	83	83	46	42	39	25	18	21	18	21	18	19	19	19
Serv.	87	88	87	84	82	81	84	44	43	40	28	24	21	21	21	18	21	21	21
Labr.	89	89	89	87	84	81	83	50	47	44	33	24	21	22	21	22	21	21	21

* Above diagonal, 1960; below diagonal, 1970.



***Key:**

- 1: white professionals
- 2: white managers
- 3: white salesmen
- 4: white clerical workers
- 5: white craftsmen
- 6: white operatives
- 7: white service workers
- 8: white laborers

- 9: nonwhite professionals
- 10: nonwhite managers
- 11: nonwhite salesmen
- 12: nonwhite clerical workers
- 13: nonwhite craftsmen
- 14: nonwhite operatives
- 15: nonwhite service workers
- 16: nonwhite laborers

Figure 1. Two-Dimensional Smallest-Space Solution for Distances Implied by Residential Dissimilarities between Employed Males Categorized by Occupation and Race, 1960*

tances as implied by intergenerational occupational mobility matrices (Blau and Duncan, 1967) and the distributions of friendship relations across categories of occupation, ethnicity, and religion (Lauermann, 1973).

The smallest-space solution for the indexes of dissimilarity calculated for 1960 is shown in Figure 1. The solution for 1970 is shown in Figure 2.⁹ In both figures we can observe along the horizontal dimension a graphic representation of the great social distance between nonwhites and whites. The segregation of occupational groups within each racial category can be observed along the vertical dimension. Both Figures 1 and 2 and the more precise metric information in Tables 3 and 4 are helpful in grasping the pattern and parameters of occupational residential segregation within and across categories of race.

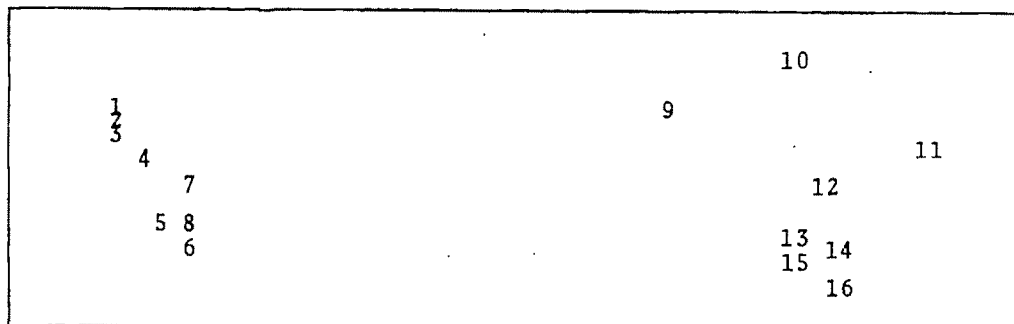
Occupational Residential Segregation Within Categories of Race

To what degree were the gross levels of occupational residential segregation due to racial residential segregation and dif-

ferences between occupations in racial composition? Comparing Table 3 with Tables 1 and 2, it is apparent that the gross levels of residential segregation between occupations were often higher than the levels of occupational segregation within either racial category. For example, in 1960 the average value of d between managers and laborers was 46 among whites and 41 among nonwhites (Table 3), while the index referring to gross segregation between managers and laborers was 59 (Table 1). Clearly, the high levels of racial residential segregation and the different racial compositions of occupational categories contributed significantly to the gross levels of occupational residential segregation. In comparisons between occupations which are similar in socioeconomic status, differences in both racial composition and residential distribution are small, and racial residential segregation contributed virtually nothing to the gross level of occupational residential segregation. However, in comparisons between occupations of very different socioeconomic status, a significant portion of the gross level of occupational residential segregation was attributable to racial segregation in occupation and residence.¹⁰

⁹ Both the axes and the scale of the smallest-space solution are arbitrary (i.e., the diagrams can be expanded or rotated with no change in the information conveyed).

¹⁰ These conclusions were also confirmed by the results of a decomposition of the gross indexes for



*For key see Fig. 1.

Figure 2. Two-Dimensional Smallest-Space Solution for Distances Implied by Residential Dissimilarities between Employed Males Categorized by Occupation and Race, 1970*

Of course, the degree to which racial segregation in both occupation and residence contributed to the gross levels of occupational residential segregation was also dependent upon the percentage nonwhite in each area at each time. Thus, comparisons between urbanized areas of gross levels of occupational residential segregation, as reported by Wilkins (1956) and Uyeki (1964), largely reflect differences between areas in racial composition.

Was occupational residential segregation as great among nonwhites as among whites? In examining Table 3, one can see that levels of occupational segregation were broadly similar for both whites and nonwhites. These findings are consistent with those reported earlier by Kantrowitz (1973) regarding segregation between persons categorized by education and income. One apparent exception to these similarities is the relatively high level of segregation among the three highest occupational categories of nonwhites. How-

ever, the number of nonwhites in these occupations was small in those urbanized areas with small nonwhite populations (such as Syracuse in 1960). The index of dissimilarity can be inflated in such circumstances due to both random deviations from an even distribution and the impossibility of distributing a small number of persons perfectly evenly over a large number of tracts.

The changes between 1960 and 1970 in residential segregation between employed men differentiated by both occupation and race are displayed in Table 4. By comparing the portions of Table 4 referring to occupational segregation within each racial category with the half of Table 2 below the diagonal, the within-race changes in occupational residential segregation between 1960 and 1970 can be compared to the gross changes. The gross changes in occupational residential segregation during the 1960s mostly reflected changes in the distribution of whites. The 1960-1970 changes among nonwhites do not show the same pattern as that seen among whites or in the gross pattern of changes. More specifically, the decrease in gross segregation between service workers and laborers vis-à-vis the higher occupational categories was apparently due to changes among whites. Such changes were not observed among nonwhites. Furthermore, while these changes among whites were fairly uniform from urbanized area to urbanized area, the changes among nonwhites were quite erratic and displayed no consistent pattern.

several cities. Winsborough's (1974) generalization of the standardization-based method of deriving components of a difference between two rates was used to decompose the gross indexes of occupational residential segregation into portions due to (1) racial segregation, (2) occupational segregation, and (3) the interaction component. In some cases, the portion due to racial segregation in residence and occupation accounted for as much as one-fourth of the gross level of occupational residential segregation. However, the size of this component varied with the urbanized area and the occupational contrasts involved.

Table 4. Change 1960-1970 in Indexes of Residential Dissimilarity Calculated between Employed Males Categorized by Occupation and Race, Averaged over the Ten Urbanized Areas

[illegible]

*Occupational Residential Segregation
between Categories of Race*

Whites and blacks have been found to be very highly residentially segregated from each other regardless of social class (Taeuber, 1968; Kantrowitz, 1973). Nevertheless, their residential distributions are not totally separate, and there may be small differences in the degree of racial segregation depending on the status of the whites and nonwhites involved. In order to grasp a fuller picture of residential distribution by social status, we now turn our attention to the ways in which differences in occupational status and racial status in combination result in varying degrees of residential segregation. The degree to which racial residential segregation differed, depending upon the respective occupational categories of the members of each race compared, is indicated in the square submatrices in Tables 3 and 4 which refer to interracial contrasts.

From Table 3 and Figures 1 and 2, it is clear that racial residential segregation in both 1960 and 1970 was indeed quite high regardless of occupation. In both Figures 1 and 2 the points representing each category of occupation by race fall into two distinct clusters, one representing whites and the other representing nonwhites. Within each racial cluster the occupational categories are arrayed roughly in line along a dimension which corresponds closely with their ranking in socioeconomic status (with the exception of the inversions noted previously).

Despite the fact that the degree of racial residential segregation was high in all instances, there were small but consistent patterns of variation in such segregation—depending upon the occupational categories involved. Furthermore, there were changes in these variations between 1960 and 1970. In 1960 whites and nonwhites in the upper occupational categories were slightly more segregated than were whites and nonwhites in the lowest occupational categories. An exception to this overall pattern involved nonwhite professionals; they were the least segregated from upper-status whites, and they were less segregated from lower-

status whites than were nonwhites in the middle occupational categories.

How did these aspects of residential segregation by race and occupation change between 1960 and 1970? In examining Table 4, it is apparent that during the 1960s nonwhites in the highest occupational categories became slightly less segregated from whites of all occupational categories. Concurrently, white and nonwhite laborers became slightly more segregated. In general, the higher the occupational status of the whites and nonwhites compared, the greater the decrease in residential segregation.

By 1970 nonwhite professionals were clearly less segregated from whites of all occupations than were other nonwhites. The differences between nonwhite professionals and other nonwhites in this regard were not very great in magnitude (on the order of 4 to 12 points) but the pattern of differences was quite consistent. Apart from the comparisons involving nonwhite professionals, nonwhites and whites in the lowest occupational categories were still slightly less segregated than those in the higher categories.

Considering the sizable expansion of black residential areas and the suburbanization of the white population which took place during the 1960s, the cross-racial dissimilarities in residential distribution remained surprisingly stable. Nevertheless, even though the changes in the indexes involving interracial comparisons were small in magnitude, the pattern of these changes is interesting. Previous studies have reported that gross racial residential segregation decreased in most U.S. cities during the 1960s (Sørensen et al., 1975). The present analysis suggests that these changes in gross racial segregation may have been due mostly to decreases between whites and upper-status blacks. Historically, blacks with high incomes have been as highly or more highly segregated from whites with similar incomes than have low-income blacks (Taeuber, 1968). Assuming income to be closely related to occupational status, this evidently became slightly less true in these ten areas by 1970. The small decreases in segregation between upper-

status nonwhites and whites between 1960 and 1970 suggest that the civil rights and open housing legislation of the 1960s may not have been totally without effect. However, the meager changes occurring during this period may have affected only those blacks whose ability or desire to become less segregated from whites was least impeded by a lack of either financial resources or occupationally-derived social prestige.

Conclusions

The main findings can be briefly reiterated.

1. During the 1950s gross occupational residential segregation in these urbanized areas increased slightly between most occupational categories. The only decreases involved segregation between clerical workers and those in the lower occupational categories. During the 1960s gross segregation between service workers and laborers vis-à-vis those in the higher occupational categories decreased, while gross segregation between the other occupational categories remained fairly constant.

2. As much as one quarter of the gross occupational residential segregation in 1960 and 1970 could be attributed to racial segregation in both the occupational and residential distributions. The degree to which racial segregation contributed to gross levels of occupational residential segregation depended upon the difference in socioeconomic status between the occupational categories compared and upon the proportion nonwhite in the urbanized area's population.

3. The 1960-1970 changes in gross occupational residential segregation were due to changes in the distribution of whites. Changes in occupational residential segregation among nonwhites showed no consistent pattern and varied greatly between urbanized areas.

4. In both 1960 and 1970, the degree of racial residential segregation was quite high, regardless of the respective occupational categories of the whites and nonwhites compared; yet there was a small but consistent pattern of variation in these

indexes. In 1960, the lowest levels of racial residential segregation were seen between whites and nonwhites in the lowest occupational categories and between nonwhite professionals and upper-status whites. During the 1960s, slight decreases were seen in racial segregation between whites and nonwhites in the highest occupational categories.

With the exception of the decrease in the segregation of laborers and service workers from persons in the higher occupational categories, perhaps the most striking aspect of the present findings is the relatively small degree of change and the absence of constant trends. It seems surprising that the substantial increases between 1950 and 1970 in the geographical sizes and population sizes of these urbanized areas did not result in substantial and consistent changes in the degrees of residential segregation. During the 1950s the population of each of these ten urbanized areas increased by an average of 25%, yet the indexes of occupational residential dissimilarity increased only slightly. During the following ten years these areas increased in population size by an average of 35%, yet most of the indexes remained the same, and the indexes computed between the highest and lowest occupational categories significantly decreased. A hypothesis that the growth of urbanized areas results in increasing residential differentiation and segregation would be inconsistent with the evidence presented here.

The residential distributions of these urbanized areas not only expanded during this twenty-year period; they also changed in their overall patterns of residential location by social status. Between 1950 and 1970, differential rates of suburbanization resulted in persons in the highest occupational, educational, and income categories being increasingly overrepresented on the peripheries of U.S. urbanized areas (Schnore, 1972). Yet in these ten urbanized areas this shifting was accompanied by only slight increases in gross occupational residential segregation in the 1950s, and by stability or decreases during the 1960s.

While persons in the highest occupa-

tional categories were redistributed toward the periphery of these urbanized areas, the residential distribution of white blue-collar workers also shifted away from the central cities. The relocation of many businesses and manufacturing concerns outside the central cities during the 1960s was accompanied by the residential relocation of many white blue-collar workers. It is possible that the degree of segregation between white blue-collar workers and whites in the higher occupational categories was smaller in the suburbs than in the central cities (Fine et al., 1971). In addition, it is likely that the relocation of blue-collar workers outside the central cities lessened occupational residential segregation among whites on the basis of central city vs. suburban residential location. If this was the case, the suburbanization of the white blue-collar population might account for the observed decreases in the residential segregation of white laborers and service workers from white-collar workers. However, one would expect similar or greater decreases in the segregation of craftsmen and operatives from white-collar workers, and such decreases were not observed.

Apparently, neither changes in the size of these urbanized areas, nor changes in the location of residential areas differentiated by the occupation of their residents were accompanied by consistent trends in occupational residential segregation. Instead, the present findings demonstrate the impressive degree to which occupational residential segregation has been maintained at a relatively constant level in the face of a considerable degree of residential redistribution. Rather than being artifacts of the historical peculiarities of each urbanized area at each point in time, the social distances manifested and constituted by such residential segregation reflect a social space whose dimensions have been constantly maintained as the geographical location of that space expanded and changed.

REFERENCES

- Blau, Peter M. and Otis Dudley Duncan
1967 *The American Occupational Structure*. New York: Wiley.
- Cortese, Charles F., R. Frank Falk, and Jack K. Cohen
1976 "Further considerations on the methodological analysis of segregation indexes." *American Sociological Review* 41:630-7.
- Duncan, Otis Dudley and Beverly Duncan
1955a "Residential distribution and occupational stratification." *American Journal of Sociology* 60:493-503.
1955b "A methodological analysis of segregation indexes." *American Sociological Review* 20:210-7.
- Fine, John, Norval D. Glenn and J. Kenneth Monts
1971 "The residential segregation of occupational groups in central cities and suburbs." *Demography* 8:91-101.
- Kantrowitz, Nathan
1973 *Ethnic and Racial Segregation in the New York Metropolis*. New York: Praeger.
- Laumann, Edward O.
1973 *Bonds of Pluralism: The Form and Substance of Urban Social Networks*. New York: Wiley.
- Lingoes, James C.
1973 *The Guttman-Lingoes Nonmetric Program Series*. Ann Arbor: Mathesis Press.
- McFarland, David D. and Daniel J. Brown
1973 "Social distance as a metric: a systematic introduction to smallest space analysis." Pp. 213-53 in Edward O. Laumann (ed.), *Bonds of Pluralism: The Form and Substance of Urban Social Networks*. New York: Wiley.
- Schnore, Leo F.
1972 *Class and Race in Cities and Suburbs*. Chicago: Markham.
- Sørensen, Annemette, Karl K. Taeuber and Leslie J. Hollingsworth, Jr.
1975 "Indexes of racial residential segregation for 109 cities in the United States: 1950 to 1970." *Sociological Focus* 8:125-42.
- Taeuber, Karl E.
1968 "The effect of income redistribution on racial residential segregation." *Urban Affairs Quarterly* 4:5-14.
- Taeuber, Karl E. and Alma F. Taeuber
1965 *Negroes in Cities*. Chicago: Aldine.
1976 "A practitioner's perspective on the index of dissimilarity." *American Sociological Review* 41:884-9.
- U.S. Bureau of the Census
1952 *Census of Population: 1950, Vol. 2*. Washington, D.C.: U.S. Government Printing Office.
1961 *Census of Population: 1960, Vol. 1*. Washington, D.C.: U.S. Government Printing Office.
1962 "Changes between the 1950 and 1960 occupation and industry classification." Technical Paper No. 18. Washington, D.C.: U.S. Government Printing Office.
1972a *1960 Census of Population and Housing Census Tract Summary Tape*.
1972b *1970 Census of Population Summary Tapes, Fourth Count, File A (Tracts)*.
1972c *1970 MED-List Tapes*.

- 1972d "1970 occupation and industry classification systems in terms of their 1960 occupation and industry elements." Technical Paper No. 26. Washington, D.C.: U.S. Government Printing Office.
- 1973 1970 Census of Population, Vol. 1. Washington, D.C.: U.S. Government Printing Office.
- Uyeki, Eugene S.
1964 "Residential distribution and stratification, 1950-1960." *American Journal of Sociology* 69:491-8.
- 1975 "Occupation and residence: Cleveland, 1940-1970." Paper presented at the Annual Meeting of the American Sociological Association, San Francisco.
- Wilkins, Arthur H.
1956 *The Residential Distribution of Occupation Groups in Eight Middle-Sized Cities of the United States in 1950*. Unpublished Ph.D. dissertation, Department of Sociology, University of Chicago.
- Winsborough, Halliman H.
1974 "A note on the decomposition of indexes of dissimilarity." Discussion Paper 201-74. Institute for Research on Poverty, University of Wisconsin, Madison.
- Winsborough, Halliman H.; Karl E. Taeuber and Annemette Sørensen
1975 "Models of change in residential segregation 1940-1970." Paper presented at the Annual Meeting of the Population Association of America, Seattle.

RACE AND INVOLVEMENT IN COMMON LAW PERSONAL CRIMES*

MICHAEL J. HINDELANG

State University of New York

American Sociological Review 1978, Vol. 43 (February):93-109

Most contemporary sociological theories of crime predict that blacks will be overrepresented among arrestees in common law personal crimes. These theories differ, however, in the extent to which this overrepresentation is attributed to disproportionate involvement in criminal offenses vs. criminal justice system selection biases. Studies that have relied upon official data have generally supported the differential involvement hypothesis, whereas studies relying on self-report techniques generally have supported the differential selection hypothesis. National victimization survey data on victims' reports of racial characteristics of offenders are introduced as a third measurement technique in order to shed additional light on this controversy. These data for rape, robbery, and assault, are generally consistent with official data on arrestees and support the differential involvement hypothesis. Some evidence of differential selection for criminal justice processing is found; however, most of the racial disproportionality in arrest data is shown by victimization survey data to be attributable to the substantially greater involvement of blacks in the common law personal crimes of rape, robbery, and assault. These results suggest that traditional admonitions against using arrest data as an index of involvement in these crimes may be overly cautious. In fact, the results imply that more caution should attend the use of self-report data in this vein and that more attention should be given to sampling and instrument concerns in self-report techniques. As currently used, the method may not be adequate for assessing the correlates of serious illegal conduct. The results also suggest that research emphasis be placed on those theories, such as the subcultural and differential opportunity perspectives, which attempt to explain differential racial involvement in these common law personal crimes.

INTRODUCTION

One of the most important theoretical questions facing criminology is whether, and to what extent, race is related to involvement in common law personal

crimes.¹ Perhaps because of the sensitivity of this question, assertions and speculations about this relationship have outdistanced the research attention focused on

* Special gratitude is due to Michael R. Gottfredson who offered critical comments on several drafts of this paper.

¹ Included among the common law crimes are murder, rape, robbery, assault, burglary, larceny, and arson; the first four of these are designated personal common law crimes in this discussion. See Blackstone (1778) and Clark and Marshall (1967).

it. Despite the limited research on the relationship between race and common law personal crime, this relationship is central to many contemporary criminological theories, and therefore it deserves the empirical attention that must be given to any important theoretical question.

Most sociological theories of crime agree that blacks are disproportionately arrested for common law personal crimes but disagree about the extent to which arrest data are indicative of disproportionate offending behavior vs. criminal justice system selection biases. Although most sociological theories of crime can accommodate a race-crime relation in arrest data, it is possible to conceive of these theories as falling along a continuum in terms of the proportion of variation in racial differences in rates of arrest that is attributed to differential *involvement* in common law crimes vs. differential *processing* by the agents of the criminal justice system. For purposes of discussion, "low" on the continuum will refer to theories which attribute very little of the variation in racial differences in rates of arrest to differences in the involvement of blacks and whites in common law crimes. "High" on the continuum will refer to theories which attribute much of the variation to actual behavioral differences between blacks and whites.²

Theories which fall at the higher end of the continuum suggest that concomitants of race, such as cultural or economic factors, are etiologically related to involvement in criminal activity. Characteristic of theories in this range on the continuum is Wolfgang and Ferracuti's (1967) subculture-of-violence thesis. In the *Subculture of Violence*, the authors observe:

Statistics on homicide and other assaultive crimes in the United States consistently show that Negroes have rates between four and ten times higher than whites. Aside from a critique of official statistics that raises serious questions about the amount of Negro crime, there is no real evidence to deny the greater involvement that Negroes have in

assaultive crimes. . . . There is reason to agree . . . that whatever may be the learned responses and social conditions contributing to criminality, persons visibly identified and socially labeled as Negroes in the United States appear to possess them in considerably higher proportions than do persons labeled white. Our subculture-of-violence thesis would, therefore, expect to find . . . [widespread] learning of, resort to, and criminal display of the violence value among minority groups such as Negroes. (264)

Thus, although Wolfgang and Ferracuti (1967) acknowledge the difficulties in using official police statistics to investigate the relationship between race and crime, they nonetheless conclude that blacks are disproportionately *involved* in assaultive crimes. More generally, the theories of Merton (1938) and Cloward and Ohlin (1960), because of their emphasis on structural impediments to the achievement of success goals via legitimate avenues, are also consistent with the hypothesis that blacks will have rates of involvement in common law criminal activity which exceed those of whites.

Theories which fall toward the middle of the continuum include those commonly referred to as conflict theories. These theories suggest that laws are enacted to protect the interests of the more powerful segments of society, and hence many of the activities that are criminalized are activities in which less powerful persons (blacks, the poor, the young, etc.) are disproportionately involved. Although such theories generally imply or state that blacks and other less powerful members of society will be disproportionately *involved* in common law criminal activity, these theories strongly emphasize that the overrepresentation of blacks in arrest statistics is considerably influenced by enforcement practices that discriminate against the less powerful.

Within the conflict group, it is perhaps Bonger (1916; 1943) who places highest on this continuum. In *Criminality and Economic Conditions*, Bonger (1916:379) suggested that in "every society which is divided into a ruling class and a class ruled, penal law has been principally constituted according to the will of the former." Under capitalism, individuals

² In a paper on minorities as victims of police shootings, Goldkamp (1977) has presented a dichotomous characterization of labeling theories vs. theories that predict greater involvement of blacks in violent crimes.

are encouraged to use any means available, including criminal activity, to obtain material goods. Furthermore, the widespread poverty that abounds under capitalism "kills the social sentiment in man, destroys in fact all relations among men" and fosters criminality (1916:436). After having studied prison statistics in the United States, he concluded that:

Crime among Negroes is significantly higher than among whites. It is three or four times higher among the men, and four or five times higher among the women. To me this appears to eliminate the idea that actual criminality among Negroes is no greater than among whites. (1943:43)

It was Bonger's (1943) view that the higher rate of crime among blacks was attributable to the unfavorable economic circumstances in which blacks were disproportionately found.

Because of a greater emphasis on differential enforcement, other conflict theorists place lower on the continuum than does Bonger. Quinney (1970:129-30), for example, notes that "the differences in arrest rates are not, however, due entirely to the fact that Negroes may be involved more than whites in law-violating behavior, but that in similar situations Negroes are more likely than whites to be apprehended." Chambliss (1969:856) and Chambliss and Seidman (1971) emphasize that the "administration of the criminal law is a highly selective process and involves the use of a wide range of discretion" that results in "systematic bias in law enforcement." Although most of their discussion of such biases is couched in terms of lower-class persons, many of their examples are of discrimination against blacks. They observe that such persons are "more likely to be scrutinized and therefore to be observed in any violation of the law and more likely to be arrested if discovered under suspicious circumstances" (Chambliss, 1969:86; see also Chambliss and Seidman, 1971:322-46). In general, the conflict theorists are careful not to deny that personal characteristics such as race are related to involvement in criminal behavior but tend instead to focus on power-related differentials in enforcement

which introduce an unspecified proportion of contamination into arrest statistics.

At the lowest point on the continuum, those who subscribe to the labeling perspective in its most extreme form are found. Those at this pole argue that there are no demographic, sociological, or psychological correlates of involvement in criminal behavior but rather that any differences in arrest rates along social, demographic, or other dimensions are attributable to biases in official processing. Characteristic of this position is the thesis presented by Chapman (1968:4) in *Sociology and the Stereotype of the Criminal*:

... 3. That apart from the factor of conviction there are no differences between criminals and non-criminals.

4. That criminal behavior is general, but the incidence of conviction is controlled in part by chance and in part by social processes which divide society into the criminal and non-criminal classes, the former corresponding to, roughly, the poor and underprivileged.

Lemert (1967:24), in his discussion of "the new deviance sociology"—Lemert (1951), Tannenbaum (1938), Kitsuse (1964), Goffman (1961), Erikson (1962), Becker (1963)—notes,

In extreme statements deviance is portrayed as little more than the result of arbitrary, fortuitous, or biased decision-making, to be understood as a sociopsychological process by which groups seek to create conditions for perpetuating established values and ways of behaving or enhancing the power of special groups.

Lemert's exception to this extreme position places him more toward the middle of our continuum, in light of his observation that "striving to validate a conception of deviance as primarily a definitional phenomenon overlooks the way in which the societal reaction varies with *objective differences in behavior, its context, and its consequences*" (1967:21, emphasis added).

This brief review of contemporary sociological theories relevant to the question of the relationship between race and crime demonstrates that although all of these theories can accommodate higher arrest rates for blacks than whites, they differ substantially in the extent to which

such racial differences are attributable to objective differences in behavior. Because each of these positions claims support from research findings, it is necessary to turn our attention to that body of research.

PRIOR RESEARCH

Official Data

Among the sources of official data on race and crime, the Federal Bureau of Investigation's *Uniform Crime Reports* (UCR) is the most extensive. In the UCR, characteristics of arrestees, including race, are published for the United States in the aggregate. These data for 1975 show that in relation to their representation in the general population (about 11%), blacks were substantially overrepresented among arrestees for murder and nonnegligent homicide (54%), forcible rape (45%), robbery (59%), aggravated assault (40%), burglary (28%), larceny and theft (31%), and motor vehicle theft (26%). For all of the remaining (Part II) offenses, blacks constituted 22% of the arrests (Kelley, 1976).

Police data were used by Mulvihill et al. (1969) to investigate the relationship between race and crime in a study published under the auspices of the President's Commission on the Causes and Prevention of Violence. Data were gathered on a probability basis from offense and arrest reports in police files in 17 large cities covering all regions of the United States. The authors report that the race of the offender was black in 72% of the criminal homicides, in 74% of the aggravated assaults, in 70% of the forcible rapes, in 85% of the armed robberies, and in 81% of the unarmed robberies (Mulvihill et al., 1969:271-83).

Wolfgang et al. (1972:Table 5.3) in their *Delinquency in a Birth Cohort* use as their criterion recorded police contacts of their cohort when the subjects were 7 through 17 years old. For the offenses of homicide, rape, robbery, and aggravated assault, the number of contacts per 1,000 cohort subjects was 139.9 for nonwhites and 9.2 for whites. For the remaining offenses (burglary, larceny, and auto theft), the number

of contacts per 1,000 cohort subjects was 476.6 for the nonwhites and 124.0 for the whites.

These official data from the UCR, the Violence Commission study, and the Wolfgang et al. (1972) cohort study are typical of comparisons of the offending of whites and blacks as measured by a variety of official data (arrest data, victim reports to the police, and police contacts). Theorists at the higher end of the continuum use such data as evidence of disproportionate involvement by blacks in the common law crimes. Theorists at the lower end of the continuum, by definition, argue that such official data are reflective of differential selection patterns, and hence they rely heavily on research data that are generated independently of the criminal justice system to support their arguments—typically data on self-reported involvement in illegal activities. To date, such studies have focused almost exclusively on juveniles.

Self-Report Measures

Relatively few self-report studies have compared the extent and nature of delinquency among whites and blacks. Three of the earliest research efforts to do so were those of Chambliss and Nagasawa (1969), Gould (1969), and Hirschi (1969). In the Chambliss and Nagasawa (1969:73) study, lower-class white, black, and Japanese high school boys responded to self-reported delinquency questionnaires; in addition, juvenile court records were canvassed to ascertain the official delinquency status of the respondents. Despite the finding that blacks had an official delinquency rate that was substantially higher than that of whites, when self-reported delinquency was used as the criterion, whites were found to have a slightly higher rate than blacks. Similarly, Gould (1969:330) studied junior high school boys in Seattle and found that although race (white vs. nonwhite) was moderately related to official delinquency in the usual direction ($\theta = .46$), race was virtually unrelated ($\theta = .07$) to self-reported delinquency on equivalent offenses.

In his study of more than 800 black and

1,300 white boys in California high schools, Hirschi (1969:Table 14) found that the former were more likely than the latter (42% vs. 18%) to have police records. When self-reported delinquency was used as the criterion, however, 49% of the blacks and 44% of the whites reported one or more delinquent acts. Once again, the racial differences for official delinquency were found to be much greater than for self-reported delinquency.

In two national studies, Gold and his colleagues (Williams and Gold, 1972; Gold and Reimer, 1975) made sex-specific comparisons of the self-reported delinquent behavior of black and white respondents. In the first study, a national probability sample of 736 white and 101 black respondents, 13 to 16 years old, was interviewed in 1967; in the second study, a national probability sample of 481 white and 67 black respondents, 13 to 16 years old, was interviewed in 1972. For both years and sex groups, whites and blacks reported involvement in 17 delinquent behaviors with similar frequencies. When the seriousness of eight items amenable to the Sellin-Wolfgang (1964) seriousness scoring procedure were tallied, the seriousness-weighted rate of self-reported delinquency was slightly greater for black males than for white males for both years.³

The research findings on racial differences in offending behavior can be summarized succinctly: studies using official measures of criminal and delinquent behavior (e.g., arrests) have repeatedly

found that blacks have markedly higher rates of arrest than do whites. However, studies using self-report measures of illegal behavior (almost exclusively illegal behavior of juveniles) have found that blacks and whites report only minimal differences on self-report inventories. This discrepancy between the self-report results and the official results requires some explanation and ultimate resolution, especially in light of the theoretical controversy. Thus, it is essential that a third source of data be brought to bear on this question.

THE PRESENT STUDY

Victimization Surveys

Victimization surveys, in which representative samples from the general population are asked to report on victimizations they may have suffered during a specific reference period, provide data on the relationship between race and common law crime that are independent of criminal justice system selection biases. In these surveys respondents are asked to tell interviewers about victimizations, regardless of whether or not they reported them to the police.

The data used here derive from a national survey of victims of crimes, undertaken by the U.S. Bureau of the Census under the sponsorship of the U.S. Department of Justice, commonly referred to as the National Crime Panel (NCP).⁴

There are two parts of the NCP—a national probability sample of households (and individuals) and a national probability sample of businesses. In a period of six months, six independent probability samples of households and businesses are interviewed (see LEAA, 1976 for a fuller description of sampling techniques). For purposes of estimating victimizations occurring during 1974 (the data to be used

³ For the 1967 results, Williams and Gold (1972:215) report Mann-Whitney U-test p-levels for racial comparisons on seriousness of .49 and .06 for females and males, respectively. For the 1972 results (Gold and Reimer, 1975: Tables 3 and 4) comparable U-tests are not reported. However, because the mean racial difference in seriousness rates for females is smaller in 1972 than in 1967 and the mean racial difference in seriousness rates for males is larger in 1972 than in 1967, we infer (in light of the 1967 U-test results) that in 1972 the racial difference for males is probably significant beyond the .06 level and the racial difference for females probably has a p-level of greater than .49. This inference depends, of course, on the assumption that the standard errors in seriousness rates are comparable across years. See Gold (1970:79) in which black-white differences in the seriousness and frequency of self-reported delinquency are not significant when SES is controlled.

⁴ A good deal of developmental work in the area of surveying victims of crime was undertaken by the U.S. Bureau of the Census and others prior to the implementation of the National Crime Panel. For the sake of brevity, this research will not be reviewed here, but the interested reader is referred to Hindelang (1976: Chaps. 2 and 3) for a summary of this research.

here), a total of approximately 80,000 housing units and other living quarters as well as approximately 17,000 businesses were selected for the sample (Law Enforcement Assistance Administration [LEAA], 1976:45).

From the 80,000 housing units selected to make victimization estimates for 1974, interviews were completed in about 65,000. The majority of the 15,000 housing units in which interviews were not obtained were found to be vacant, demolished, converted to nonresidential use, or otherwise ineligible to be sampled (LEAA, 1976:46). Overall, "interviews were obtained in about 96% of all eligible housing units, and about 99% of the occupants of these households participated in the survey" (LEAA, 1976:46).

In the survey of households there are three types of respondents: household respondents, self-respondents, and proxy respondents. The household respondent, usually the head of household, answers such questions as whether the residence is owned or rented and what is the family income. In addition, the household respondent answers questions about victimizations affecting the entire household, such as burglary and vehicle theft.

Self-respondents are all household respondents 14 years of age or older. In addition to background information (e.g., age, sex, and education), each of these household members is asked a series of "screen" questions designed to elicit whether or not the person has been the victim of a crime of rape, robbery, assault, or personal larceny during the preceding six months. After all screen questions have been asked, a series of detailed incident questions is asked about each victimization uncovered in the screen questions. Proxy respondents are used to elicit the same information about household members 12 and 13 years old and for individuals who are physically or mentally unable to answer for themselves. Thus, the estimates derived in the household portion of the survey are national estimates for those 12 years of age or older.⁵

⁵ Some personal victimizations are classified as series victimizations and an incident report is completed only for the last crime in a series of crimes. To

In the commercial portion of the survey, the owner, manager, or someone knowledgeable about the affairs of the business is interviewed about robbery and burglary victimizations occurring during the reference period. As in the household portion of the survey, the screen method is used to elicit victimizations and is followed by a detailed incident report. Businesses eligible for interviews, but in which interviews were not completed, amounted to fewer than 1% of those eligible (LEAA, 1976:60).

The data derived from the interviews are then weighted to give estimates for the nation (see LEAA, 1976: App. 1 and 2 for a detailed discussion of the weighting procedures and standard errors). The estimates used here are for incidents rather than victimizations. In the former, one incident is counted for each event uncovered during the survey regardless of the number of victims involved in the crime, whereas in the latter, one victimization is counted for each person victimized during an event. The incident weight assigned to each sample case is adjusted in those cases in which an event involved more than one victim because such cases would have more than a single chance of being included in the sample

be counted as a series victimization *all* of the following criteria must be met:

- (a) the respondent must be unable to recall the details of the victimizations in the series well enough to report on the circumstances of each victimization separately;
- (b) the victimizations must be of a similar type; and
- (c) there must be at least three victimizations to constitute a series.

Series victimizations are tabulated separately from nonseries victimizations and tend disproportionately to be either assaults, more likely simple than aggravated, or household larcenies under \$50 (LEAA, 1976:51). I have excluded series victimizations from the analyses for two reasons: (1) because detailed information, such as the perceived race of the offender is collected only for the last crime in a series, details of series crimes are not known; and (2) because the respondent may not have reported accurately on the time of occurrence, it is less certain that these events fall within the reference period. In 1973, about 5% of the personal victimizations reported to interviewers were classified as series victimizations. See Hindelang (1976: App. F) for a discussion of some of the problems inherent in the use of series data.

(LEAA, 1976:49-50). The mean incident weight for the crimes studied here (rape, robbery, aggravated assault, and simple assault) in the household portion of the survey is about 1,023.

Because the UCR arrest data include business robbery arrestees in tables on the race of offenders, it was necessary also to include business robberies in the NCP data. This was accomplished by combining the results on the race of offender in business incidents from the 1974 NCP with those for personal robberies.

Results

In the course of the interview, respondents who had been confronted by offenders were asked a series of questions about the offender(s), including the number of offenders involved and the sex, race, and estimated age of the offender. The data presented here have been tabulated according to the responses to these questions. Before proceeding to the findings, it is necessary to discuss briefly how they were generated.

Each incident reported to survey interviewers was weighted by the number of offenders that the victim reported was involved in the incident.⁶ Incidents in which the number of offenders was unknown or not ascertained or in which there was a group of offenders of mixed races (i.e., in which some were white and some were black or of other races) were excluded from analysis. It was necessary to exclude incidents in which the number of offenders was unknown because in such cases the victim was not asked the race of the offender(s). It was necessary to exclude incidents involving multiple offenders of mixed races because victims were not asked how many offenders were from each racial group. Of the total estimated number (8,130,059) of incidents of robbery, rape, and assault elicited by the survey for the 1974 calendar year, 3% or

210,824 were excluded from the analysis because they met at least one of the exclusionary conditions noted above.

The analyses presented here are limited to robbery, rape, and assault because it is in victimizations of these types that the victim is confronted by the offender and hence is able to report on the offender's characteristics. In the crimes of burglary, household larceny, and vehicle theft, the victim usually does not see the offender during the commission of the crime and therefore cannot report on the offender's personal characteristics.

The UCR annually publishes data on the racial characteristics of arrestees. It is possible to compare the racial characteristics of offenders as reported by victims in the 1974 national sample of the National Crime Panel to the racial characteristics of arrestees as reported in the 1974 UCR. If there are substantial biases in the UCR data for *any* reason, we would expect, to the extent that victimization survey reports are unbiased, to find large discrepancies between the UCR arrest data and victimization survey reports on racial characteristics of offenders. Specifically, the theories at the lowest point of the continuum predict a very substantial overrepresentation of blacks in the population of UCR arrestees, whereas theories at the highest point of the continuum predict a very small overrepresentation of blacks in arrest statistics *when compared with victimization survey reports*.⁷ Furthermore, the former theories (low on continuum) would predict a small overrepresentation of black offenders in victimization survey reports relative to the representation of blacks in the general population, whereas the latter (high on continuum) would predict a very substantial overrepresentation. The victimization survey reports are thus an independent measure of the involvement of whites and blacks in rape, robbery, and assault—a measure that cannot be affected by the kind of contamination with which the-

⁶ As noted above, each incident had a weight that was inversely proportional to the probability of appearing in the sample. It was actually this incident weight that was weighted by the number of offenders that the victim reported having been involved in the incident.

⁷ It should be noted here that this comparison will be sensitive not only to biases on the part of police but also to such potential biases as would be introduced if white victims were more likely to notify the police when victimized by black offenders.

Table 1. National Comparisons between Uniform Crime Reports (UCR) and National Crime Panel (NCP) Estimates* for Race of Arrestees and Offenders, 1974

		White	Black	Other	Total
Rape	NCP	60% (125,890)	39% (82,873)	1% (1,847)	210,609
	UCR	49% (7,665)	48% (7,482)	3% (453)	15,600
Robbery	NCP	34% (797,246)	62% (1,465,838)	4% (105,587)	2,368,671
	UCR	35% (31,477)	62% (55,728)	3% (2,210)	89,415
Aggravated Assault	NCP	66% (1,473,341)	30% (659,814)	4% (93,044)	2,226,199
	UCR	56% (75,136)	41% (54,870)	2% (3,330)	133,336
Simple Assault	NCP	66% (2,204,576)	29% (969,432)	5% (150,572)	3,324,580
	UCR	61% (154,757)	37% (92,417)	2% (6,337)	253,511
Representation in General Population		88%	11%	1%	

* Weighted by number of offenders in incident.

Source: UCR data, Kelley (1975:191). General population data, U.S. Bureau of the Census (1975a).

orists falling at the lower end of the continuum contend police data are rife.⁸

The UCR victimization comparisons for 1974 are presented in Table 1. Both sources follow the U.S. Bureau of the Census convention of counting Spanish-Americans as white; "other races" include American Indian, Chinese, and Japanese. As expected, for each of the offenses shown, the ratio of offenders in incidents reported by victims in the survey to the number of UCR arrestees is large. This results primarily from three factors: first, the victimization survey data include many crimes that have not been reported to the police; second, even for reported crimes, the clearance rates for these offenses are small;⁹ third, the UCR data on race of arrestees in 1974 were based on reports of police agencies covering an estimated population of 124 million persons rather than on the entire population (211 million), whereas the victimization survey results are estimates of victimization experiences of the entire U.S. population 12 years of age or older.

⁸ Some of the limitations of victim survey data are discussed later.

⁹ In 1974, 51% of the forcible rapes, 63% of the aggravated assaults, and 27% of the robberies were cleared by arrest (Kelley, 1975:43).

For convenience, the discussion will focus on the results for blacks. Table 1 shows that the UCR and the victimization data are identical for the crimes of robbery: 62% of the victimization survey offenders and 62% of the UCR arrestees were reported to have been black. For the remaining crime categories (rape and assault), in relation to the victimization reports, blacks are overrepresented by about ten percentage points in the UCR arrest data. Thus, for the crimes of rape and assault, but not for robbery, these results are consistent with the hypothesis that a small proportion of the white/black discrepancy in arrest rates is attributable to selection bias of some sort. The nature of this selection bias and some competing hypotheses will be discussed below.

In 1974 an estimated 88% of the U.S. population was white, 11% was black, and 1% was of other races (U.S. Bureau of the Census, 1975a:26). Regardless of whether the UCR or the victimization survey data are taken as the indicator, blacks are substantially overrepresented in relation to their representation in the general population. This overrepresentation is by a factor of three or three and one-half times for assault, about four times for rape, and more than five times for robbery. In light of these data, it is difficult to argue that

blacks are no more likely than whites to be *involved* in the common law crimes of robbery, forcible rape, and assault. At the same time, in UCR arrest data for assault and rape, blacks are found with a greater relative frequency than victimization survey data would predict, under the assumption that race-linked biases have not been introduced at some point in the arrest selection process.

Both the NCP and the UCR data can be dichotomized into groups of offenders (or, in the case of the UCR, arrestees) under 18 years of age and 18 years of age or older. Although not shown in tabular form, these results closely parallel those reported in Table 1.¹⁰ Thus, both among adults and juveniles, blacks are substantially overrepresented in relation to their representation in the general population regardless of whether the NCP data or the UCR data are used.

One major source of potential bias in the UCR data, one that is independent of the actions of criminal justice functionaries, is differential reporting of crimes by victims to the police. It is possible that the UCR/NCP differences, as reported in Table 1, are attributable not to discriminatory enforcement patterns but rather are attributable to victims' reporting to the police offenses committed by blacks proportionately more often than offenses committed by whites. Because research has demonstrated that offenses of these types almost exclusively come to the attention of the police through victim reporting (Reiss, 1971; Hindelang and Gottfredson, 1976), the reporting decision obviously affects arrest patterns; offenses not reported to the police almost certainly will not result in an arrest.¹¹

The data in Table 2 indicate that when only those NCP victimizations which the

victims told interviewers were reported to the police are considered, blacks are found to constitute 47% of the NCP rape offenders as compared with 48% of the UCR rape arrestees (Table 1). For both aggravated and simple assault, on the other hand, there is a slightly greater discrepancy between the NCP and the UCR percentages when only those NCP crimes which victims said were reported to the police are considered. For instance, 26% of the offenders involved in aggravated assaults which victims said were reported to the police were black, whereas 41% of the aggravated assault arrestees were black. If we operationally define criminal justice system selection bias as the discrepancy between the proportion of blacks involved in victimizations which victims said were reported to the police and the proportion of blacks among UCR arrestees, then these data suggest that there is virtually no criminal justice system selection bias for either rape or robbery but that there is such bias for assault, especially aggravated assault. Thus, once the victim's reporting (to the police) behavior is taken into account, differences between NCP crimes reported to the police and UCR data on arrestees remain only for assault.

Some NCP Measurement Problems

It is essential to note at this point that the crime of assault is the NCP crime which has the most measurement difficulties. In reverse record checks (U.S. Bureau of the Census, 1970a; 1970b; LEAA, 1972), checks in which victims are selected from police files and ideally interviewed on a double-blind basis by Bureau of the Census interviewers, assault has been found to be the most poorly measured offense in the sense that: (a) assault victims from police files are more difficult, for a variety of reasons, than other victims to locate and interview; (b) of those victims who are interviewed, a smaller proportion of assault victims than other victims from police files mention having been victimized when they are interviewed.

In the three reverse record check studies, assault victims selected from

¹⁰ For the UCR data, see Kelley (1975:192-3). Forcible rapes committed by offenders under 18 years of age are relatively rare; the number of NCP sample (unweighted) cases is very small (23 cases).

¹¹ In a small proportion of victimizations (about 3% of the victimizations in the NCP data), the police do come across the crime in progress. In addition, it is possible for police to discover incriminating evidence in the possession of an offender (e.g., another person's wallet) even though the victim may not have reported it to the police.

Table 2. Race of Offender in National Crime Panel Estimates^a (1974) by Whether the Crime Was Reported to the Police

	White	Black	Other	Total
Rape				
Not Reported to Police	66% ^b 62% ^c (78,058)	33% 48% (39,442)	1% 68% (1,261)	100% 56% (118,762)
Reported to Police	52% 38% (47,832)	47% 52% (43,431)	1% 32% (585)	100% 44% (91,848)
Total	60% 100% (125,890)	39% 100% (82,873)	1% 100% (1,847)	100% 100% (210,609)
Robbery				
Not Reported to Police	38% 47% (370,800)	57% 38% (550,378)	5% 42% (44,405)	100% 41% (965,583)
Reported to Police	30% 53% (426,445)	65% 62% (915,454)	4% 58% (61,182)	100% 59% (1,403,081)
Total	34% 100% (797,245)	62% 100% (1,465,832)	4% 100% (105,587)	100% 100% (2,368,664)
Aggravated Assault				
Not Reported to Police	62% 43% (629,837)	33% 52% (341,198)	5% 53% (49,123)	100% 46% (1,020,158)
Reported to Police	70% 57% (843,503)	26% 48% (318,617)	4% 47% (43,921)	100% 54% (1,206,041)
Total	66% 100% (1,473,341)	30% 100% (659,814)	4% 100% (93,044)	100% 100% (2,226,199)
Simple Assault				
Not Reported to Police	65% 64% (1,401,374)	31% 69% (669,775)	4% 64% (95,904)	100% 65% (2,167,053)
Reported to Police	69% 36% (803,202)	26% 31% (299,657)	5% 36% (54,668)	100% 35% (1,157,527)
Total	66% 100% (2,204,576)	29% 100% (969,432)	5% 100% (150,572)	100% 100% (3,324,580)

^a Weighted by number of offenders in incident.^b Row percent.^c Column percent.

police files had the lowest rate of completed interviews, about three out of five. Furthermore, among those victims selected from police files with whom interviews were completed, assault victims consistently had the poorest "recall" rate, that is, the smallest proportion of known victims who reported to survey interviewers that they had been victims of the crime selected from police files. For example, in completed interviews in all three studies combined, 88% of the burglary victims,

80% of the robbery victims, and 67% of the rape victims reported the crime selected from police files to survey interviewers; however, only 47% of the assault victims did so. In addition, it was found that when the offender was known to the victim, especially when the offender was related to the victim, the rate of "recall" was smaller than when the offender was a stranger (LEAA, 1972). This finding is important for two reasons. First, among the crimes studied here, assault is the

most likely to involve nonstrangers.¹² Second, victimization data suggest that nonstranger victimizations account for a larger proportion of black than of white victimizations.¹³ For example, in the NCP data used here, 32% of the white victims and 48% of the black victims were reportedly victimized in aggravated assaults by nonstrangers.¹⁴ If we assume that the "recall" problem for assaults in victimization surveys is comparable for nonstranger victimizations suffered by both blacks and whites, it follows that the *total* number of assaults suffered by black victims will be underestimated in victimization data more so than those suffered by whites because blacks apparently are victimized by nonstrangers in a greater proportion of the assaults than are whites. In light of this and because the victims in assaults are likely to be victimized by persons of the same race (Hindelang, 1976:184-6), we can infer that the victimization data on assault probably underestimate the proportion of black offenders.

SUMMARY AND CONCLUSIONS

These results demonstrate that both the victimization data and the official arrest data show blacks to be substantially over-represented, in relation to their representation in the general population, as offenders/arrestees for the common law crimes of forcible rape, robbery, and assault. Both data sources show that this disproportionality is greatest for robbery, followed by rape, aggravated assault, and simple assault. Furthermore, both sources of data show virtually the same percent figure for robbery but for the remaining three crimes the official data show a somewhat greater proportion of black offenders than do the victimization data. The discrepancy for rape is accounted for

by the finding that rapes involving black offenders are more likely to be reported to the police than are rapes involving white offenders. When only those rape victimizations that are reported to the police are examined, there is virtually no difference between the two sources. Parenthetically, it should be noted that the UCR/victimization survey similarity for robbery maintains when only victimizations reported to the police are studied. For both aggravated and simple assault, however, the NCP/UCR discrepancy is slightly greater when only victimizations reported to the police are examined. For the reasons noted above, the NCP data on assault must be viewed with less confidence than the NCP data for robbery and rape.

In general, the results on the disproportionate involvement of blacks in rape, robbery, and assault as shown in these nationwide victimization data are much more congruent with studies that have used police data than with studies that have used self-reports of offenders. This is true even when the data for offenders reported by victims to be under 18 years of age, the age group most often studied in self-report research, are examined separately.

There are several reasons why the self-report studies discussed earlier may be incompatible with the NCP and UCR data.¹⁵ First, with the exception of Gould's (1969) study, these self-report studies did not examine official and self-reported offenses of comparable seriousness; the self-report studies are weighted toward the least serious offenses. Second, with the exception of the Gold (1970; Gold and Reimer, 1975; Williams and Gold, 1972) studies, the self-report studies drew samples from in-school populations; the higher dropout rates for blacks than for whites means that this sampling approach results in a race-linked sampling bias. Third, the Gold studies, which increasingly are the studies most often cited by those arguing that black/white differences

¹² This holds generally whether victimization survey data or police data on offenses known are used. See Hindelang (1976: Table 7.16).

¹³ Of course, the victim-offender relationship affects the data, because, as noted above, the "recall" rate in victimization data is related to the relationship between the victim and the offender.

¹⁴ Nonstrangers include family members and other relatives, persons well-known to the victim, and casual acquaintances of the victim.

¹⁵ In the course of discussing theoretical perspectives above, some of the shortcomings of official data were suggested by the theorists. For critiques of police data see Doleschal and Wilkins (1972) and Wolfgang (1963).

are minimal, had numbers of blacks that are certainly too small for reliable conclusions in a national survey (1967: 53 black males, 48 black females; 1972: 33 black males, 34 black females), especially in light of the relative rarity of all but the most trivial illegal behaviors.¹⁶

These findings may indicate that at its current level of methodological development, the self-report technique is simply inadequate for assessing racial differences in rates of involvement in serious offenses. Even if the self-report technique itself were well developed from a methodological point of view, it would be necessary to use sample sizes much larger than those used in self-report studies to date in order to estimate reliably correlates of involvement in serious offenses.

Earlier, I discussed a continuum along which selected criminological theories fall in terms of the proportion of variation in arrest statistics which they attribute to involvement in criminal activity. It was noted that because all of these theories can accommodate the higher black arrest rate for common law personal crimes, it is necessary empirically to disentangle involvement in these crimes from selection biases in criminal justice processing. To the extent that a small percentage of the racial disproportionality in arrest rates is attributable to racial disproportionality in offending behavior, theories at the lower end of the continuum (e.g., labeling) are supported and theories at the higher end of the continuum (e.g., subculture of violence) are not. If the victimization data as reported in Table 1 are taken as a measure of involvement in rape, robbery, and assault and the UCR arrest data are taken as a measure of involvement in these crimes *plus* selection biases,¹⁷ then the discrep-

ancy between the two data sources can be taken as a measure of selection bias. The ratio of the percentage of black offenders in the NCP data to the percentage of black arrestees in the UCR data is an index of the proportion of the arrest percentage that can be attributed to criminal involvement; the complement can be attributed to selection bias. For robbery this ratio ($62\% \div 62\%$) is 1.0, indicating that none of the arrest percentage can be attributed to selection bias. The ratios for the remaining offenses are .84 for rape, .78 for simple assault, and .73 for aggravated assault. This indicates that for these crimes some of the arrest percentage can be attributed to selection bias but, by far, most of the arrest percentage appears to be attributable to the substantially greater involvement of blacks than whites in these crimes.

The complements of these ratios can be thought of as an index of selection bias. These selection biases can be decomposed into criminal justice system biases (differential police patrols, closer scrutinizing of blacks by police, greater propensity of police to arrest blacks once contact has been made, etc.) and victim reporting (to the police) biases. By using the data in Tables 1 and 2, the ratios discussed above can be recomputed, using in the numerator only those incidents which vic-

¹⁶ For example, the mean number of burglaries reported during the three years prior to the interview was .06 for the white males and .09 for the black males in the 1972 study. Also in 1972, more than one-quarter of the males had total Sellin-Wolfgang (1964) seriousness scores of zero. See Nettler (1974) and Reiss (1975) for some recent discussions of suggested methodological shortcomings of the self-report method.

¹⁷ These biases include those introduced by differential reporting of offenses by victims to the police as well as criminal justice system selection biases. If

should be noted here that the UCR data may also reflect statistical biases due to the fact that not all police jurisdictions report arrest data to the FBI. For data on offenses, the UCR received reports from jurisdictions covering 94% of the U.S. population in 1974. The published UCR arrest data are broken out only for cities, suburban areas, and rural areas. Because the offense data cover almost all of the U.S. population, we can examine the extent to which the arrest data come disproportionately from city, suburban, and rural areas by examining the ratio of the U.S. population covered by agencies that report arrest data (Kelley, 1975: Tables 44, 49 and 54) to the U.S. population covered by agencies reporting offense data (Kelley, 1975: Table 13). These ratios are .72 for the cities, .72 for the suburban areas, and .51 for the rural areas. Thus, rural areas are underrepresented in the national arrest data. However, the total number of arrests in rural areas for rape, robbery, and assault (26,000) is relatively small in relation to the total number of arrests in all reporting areas for these crimes (490,000), and even if adjustments were made for the underrepresentation of rural areas, the effect on the aggregate figures would not be substantial.

tims said were reported to the police. For example, in rapes elicited in the NCP survey which victims said were reported to the police, offenders were black in 47% of the cases (Table 2). In the UCR arrest data, blacks constituted 48% of the rape arrestees. This ratio for rape is $(47\% \div 48\%) .98$. The comparable ratio is 1.05 for robbery, .63 for aggravated assault, and .70 for simple assault. The complements of the ratios reported in the previous paragraph can be taken as a measure of all selection biases, that is, criminal justice system and victim reporting biases. The complements of the ratios reported in this paragraph can be taken as a measure of selection biases exclusive of victim reporting biases. For aggravated assault the first complement is .22 and the second is .37, which suggests that criminal justice system biases are more pronounced than they first appear once victim reporting biases are taken into account. For rape, the first complement is .16 and the second is .02, which suggests that for this crime category almost all $(.16 - .02 = .14)$ of the deviation of the UCR from NCP data is attributable to the reporting biases of victims rather than to criminal justice system selection bias.

These results suggest that criminal justice system selection bias may be greatest for aggravated and simple assault and may be negligible for rape and robbery. At the same time, these data indicate that such biases account for much less of the racial disproportionality for these crimes than does differential involvement. These results do not support the heavy theoretical emphasis on differential selection manifested in the theories at the lower end of the continuum.¹⁸ Rather, these data suggest that theories of criminality must

give more attention to explaining disproportionate involvement (or noninvolvement) in common law personal crimes among blacks and whites, as do those theories higher on the continuum.

Of course, there are some measurement problems that may affect the victimization survey results. For example, victims' reports of the racial characteristics of offenders may be affected by popular stereotypes of the criminal. Furthermore, persons of Spanish heritage may be reported by some victims to be black. Because Spanish heritage persons in the general population are counted as white by U.S. Bureau of Census convention, this potential definitional difference may artificially inflate the disproportionality of blacks among the NCP offenders. It is also possible that white victims, who in assault and rape are likely to be victimized by white offenders, underreport their victimizations to survey interviewers. On the other hand, as noted in the discussion above, available evidence indicates that race-linked biases in the measurement of assault in victimization surveys may tend to undercount black offenders. Furthermore, there is some evidence that the most undersampled respondent group consists of young black males (U.S. Bureau of the Census, 1975b), who would be expected to be victimized in these crimes by black offenders. In any event, it is unlikely that biases in the victimization surveys linked to the race of the offender would be of such a magnitude that the substantial overrepresentation of blacks in the offender population would disappear.

The results of this study have important theoretical implications. The analysis indicates that theories at the lower end of the continuum are incapable of accounting for these victimization survey results which show a much higher rate of involvement for blacks than for whites in these common law personal crimes; the data from two independent sources (victimization surveys and arrest data) are in close agreement with each other and in sharp disagreement with the predictions of labeling theory, particularly in its extreme form (e.g., Chapman, 1968). As one moves from the lower to the higher end of the continuum, the predictions of the

¹⁸ It should be stressed that I am not arguing that these results by any means indicate that racial discrimination in the criminal justice system should cease to be a cause for concern. There is evidence within the data of racially discriminatory enforcement and, obviously, any racial discrimination in the mechanisms by which people enter the criminal justice system is objectionable and demands attention. The argument being made is that it appears that these data seriously question sociological explanations which attribute most of the racial disproportionality in arrest data to differential selection rather than to differential involvement.

theories become increasingly compatible with the results. Conflict theory, which is nearer to the middle of the continuum than labeling theory, is somewhat compatible with the results in that it predicts racial differential in involvement in common law personal crimes because the more powerful (white) segments of society will have legislated against those activities in which the less powerful (black) segments of society engage disproportionately. However, many conflict theorists (e.g., Chambliss, 1969; Chambliss and Seidman, 1971; Quinney, 1970) also strongly emphasize that power differentials result in differential processing by agents of the criminal justice system. Although there was evidence of bias in police processing for assault, there was virtually none for robbery and rape.

Theories at the higher end of the continuum are most consistent with the data. These theories, such as Merton's (1938) anomie theory, Cloward and Ohlin's (1960) opportunity theory, or Wolfgang and Ferracuti's (1967) subculture-of-violence theory are quite compatible with large racial differentials in involvement in common law personal crimes and comparatively small racial differentials in police processing.

It is much more difficult, in light of the findings, to choose among theories at the higher end of the continuum than it is to reject, as incompatible with the data, theories at the lowest end. For example, these findings could easily be interpreted within the differential opportunity perspective because the NCP and UCR data indicate that the greatest racial differences in involvement are for the economically motivated crime of robbery.¹⁹ Blocked

access to legitimate avenues for material achievement (e.g., Merton, 1938; Cloward and Ohlin, 1960) can be readily invoked to account for the higher rate of black involvement in robbery. Although differential opportunity might seem less capable of encompassing the higher rates of black assault and rape, several strain theorists have suggested mechanisms to account for violent crime. Cloward and Ohlin (1960:171-8) postulate the existence of a subcultural adaptation to blocked opportunities that is organized around violent behavior. In this conflict subculture, "... violence is the keynote; its members pursue status ('rep') through the manipulation of force or threat of force" (1960:20). Under conditions conducive to the development of conflict subculture:

... tendencies toward aberrant behavior become intensified and magnified. These adolescents seize upon the manipulation of violence as a route to status not only because it provides a way of expressing pent-up anger and frustrations but also because they are not cut off from access to violence by vicissitudes of birth. In the world of violence, such attributes as race, socioeconomic position, age, and the like are irrelevant. . . . (1960:175)

Similarly, Cohen (1955) suggests a strain model which links school failure with an increased probability of engaging in negativistic, nonutilitarian, and malicious behavior. One important emphasis of the anomie theorists which cannot be ignored in accounting for differential rates of violent behavior is their emphasis on the frustration that accompanies blocked opportunities and the use of violence as status-conferring.

On the other hand, those subscribing to the subculture-of-violence perspective have argued that blacks are more likely than whites to be members of the violent subculture and hence are more likely than whites to accept and expect violent behavior in social interactions (Wolfgang and Ferracuti, 1967; Curtis, 1975). To the extent that the crimes studied here, that is, rape, robbery, and assault, are construed

¹⁹ Another NCP property crime, personal larceny, also shows greater racial differences than those found for assault and rape. A small proportion of larcenies involve a confrontation between the victim and the offender and are differentiated from robberies in that they do not involve force or the threat of force—for example, a purse snatch in which force is not threatened or directed at the victim. The UCR arrest tables do not separate larcenies of this type from larcenies in which there is no personal confrontation between the victim and the offender; the NCP data on larcenies that do involve such a confrontation cannot reasonably be compared with the UCR

arrest data. However, the 1974 NCP data reveal that 69% (estimated N = 230,091) of the offenders in face-to-face larcenies were black.

to be violent crimes, the results are interpretable within this subcultural perspective. It should be noted, however, that despite the fact that robbery has been defined by some as a violent crime (e.g., in the Uniform Crime Reports; Mulvihill et al., 1969), the primary aim of robbery is to deprive a person of property, whereas assault and rape, if completed, necessarily involve bodily harm to the victim. Among the personal crimes studied here, the least violent crime (robbery) shows the greatest racial difference,²⁰ and the more violent crimes (assault and rape) show a smaller racial difference. Subculture-of-violence theorists argue, however, that much violent crime is intraracial, particularly blacks victimizing blacks. It may be that intraracial crimes of assault and rape are undercounted in both the NCP and UCR data because, according to this subcultural theory, they are more accepted and expected by blacks in social interactions and hence are less often construed as crimes and/or reported to either the police or to survey interviewers as crimes. Of course, the resolution of this speculation is well beyond the scope of the data.

By the definition of the continuum used here, the theories at the higher end all predict racial differences in involvement in common law personal crimes and the data are not sufficient for choosing among these competing explanations. However, the results do strongly indicate that research attention in this area should focus on these competing (higher-end) explanations of the differential involvement of blacks and whites in common law personal crimes. Is the differential attributable to the disparity between the socioeconomic status distributions of blacks and whites as Bonger (1943) has suggested or to actual or perceived differences in structural impediments to the achievement of success goals via legitimate avenues (e.g., Merton, 1938; Cloward and Ohlin, 1960)? Are these differences due to cultural factors as Wolfgang and Ferracuti (1967) argue? Or is the explanation to be found in the historical maltreatment of American blacks? Only further research specifically designed to assess these and other hy-

potheses can shed additional light on these questions.

Throughout this paper, substantial care has been taken to limit the discussion to common law crimes of forcible rape, robbery, and assault. The NCP data now available only address common law crimes. For the crimes of burglary, household larceny, and vehicle theft, which are also included in the NCP surveys, offender characteristics are generally unavailable because these crimes do not typically involve a face-to-face confrontation between the victim and the offender. Clearly, these results cannot be extrapolated beyond the specific crimes to which the analyses were addressed. If the differential involvement in white-collar offenses, organized crime, corporate crime, or consumer fraud had been studied, the results might have been very different. Obviously, these data and analyses shed no light on racial differences in crime generally.

As Nettler (1974:126) has noted:

... caution is required in the interpretation of differentials in crime rates between whites and nonwhites. In the light of the sad history of racial relations, it is difficult to make comparisons today of the relative importance of the alleged causes of any differences in observed behaviors.

Although research on this question is difficult for a variety of reasons, the results of this study suggest that it is incumbent upon social scientists to give long overdue research attention to such basic questions as these.

REFERENCES

- Becker, Howard S.
1963 *The Outsiders: Studies in the Sociology of Deviance*. New York: Free Press.
- Blackstone, William
1778 *Commentaries on the Laws of England*. 8th ed. Oxford: Clarendon Press.
- Bonger, Willem
1916 *Criminality and Economic Conditions*. Tr. Henry Horton. Boston: Little, Brown.
- 1943 *Race and Crime*. Tr. M. Hordyk. Montclair: Patterson-Smith.
- Chambliss, William
1969 *Crime and the Legal Process*. New York: McGraw-Hill.
- Chambliss, William and Richard Nagasawa
1969 "On the validity of official statistics: a

²⁰ Similarly for personal larceny (see fn. 19).

- comparative study of white, black, and Japanese high school boys." *Journal of Research in Crime and Delinquency* 6:71-7.
- Chambliss, William and Robert Seidman
1971 *Law, Order, and Power*. Reading, Pa.: Addison-Wesley.
- Chapman, Dennis
1968 *Sociology and the Stereotype of the Criminal*. London: Tavistock.
- Clark and Marshall
1967 *A Treatise on the Law of Crimes*. 7th ed. New York: Holt.
- Cloward, Richard and Lloyd Ohlin
1960 *Delinquency and Opportunity: A Theory of Delinquent Gangs*. New York: Free Press.
- Cohen, Albert
1955 *Delinquent Boys*. New York: Free Press.
- Curtis, Lynn
1975 *Violence, Race, and Culture*. Lexington: Lexington Books.
- Doleschal, E. and Leslie Wilkins
1972 *Criminal Statistics*. Rockville, Md.: Center for Studies in Crime and Delinquency.
- Erikson, Kai
1962 "Notes on the sociology of deviance." *Social Problems* 9:307-14.
- Goffman, Erving
1961 *Asylums*. New York: Anchor.
- Gold, Martin
1970 *Delinquent Behavior in an American City*. Belmont, Ca.: Brooks/Cole.
- Gold, Martin and David Reimer
1975 "Changing patterns of delinquent behavior among Americans 13 through 16 years old." *Crime and Delinquency Literature* 7:483-517.
- Goldkamp, John
1977 "Minorities as victims of police shootings: interpretations of racial disproportionality and police use of deadly force." *Justice System Journal* 2:169-83.
- Gould, Leroy
1969 "Who defines delinquency: a comparison of self-reported and officially-reported indices of delinquency for three racial groups." *Social Problems* 16:325-36.
- Hindelang, Michael
1976 *Criminal Victimization in Eight American Cities: A Descriptive Analysis of Common Theft and Assault*. Cambridge, Ma.: Balinger.
- Hindelang, Michael and Michael Gottfredson
1976 "The victim's decision not to invoke the criminal process." Pp. 57-78 in W. McDonald (ed.), *The Victim and the Criminal Justice System*. Beverly Hills: Sage.
- Hirschi, Travis
1969 *Causes of Delinquency*. Berkeley: University of California Press.
- Kelley, Clarence
1975 *Crime in the United States*. Washington, D.C.: U.S. Government Printing Office.
- 1976 *Crime in the United States*. Washington, D.C.: U.S. Government Printing Office.
- Kitsuse, John
1964 "Societal reaction to deviant behavior: problems in theory and method." Pp. 87-102 in Howard Becker (ed.), *The Other Side: Perspectives on Deviance*. New York: Free Press.
- Law Enforcement Assistance Administration, U.S. Department of Justice, National Institute of Law Enforcement and Criminal Justice, Statistics Division
1972 *San Jose Methods Test of Known Crime Victims*. Statistics Technical Report No. 1. Washington, D.C.: U.S. Government Printing Office.
- 1976 *Criminal Victimization in the United States*. Washington, D.C.: U.S. Government Printing Office.
- Lemert, Edwin
1951 *Social Pathology*. New York: McGraw-Hill.
- 1967 *Human Deviance, Social Problems and Social Control*. Englewood Cliffs: Prentice-Hall.
- Merton, Robert
1938 "Social structure and anomie." *American Sociological Review* 3:672-82.
- Mulvihill, D., M. Tumin and L. Curtis
1969 *A Staff Report Submitted to the National Commission on the Causes and Prevention of Violence*. Crimes of Violence, Vol. 11. Washington, D.C.: U.S. Government Printing Office.
- Nettler, G.
1974 *Explaining Crime*. New York: McGraw-Hill.
- Quinney, Richard
1970 *The Social Reality of Crime*. Boston: Little, Brown.
- Reiss, Albert Jr.
1971 *The Police and the Public*. New Haven: Yale University Press.
- 1975 "Inappropriate theory and inadequate methods as policy plagues: self-reported delinquency and the law." Pp. 211-22 in N.J. Demerath, Jr., Otto Larson, and Carl F. Schuessler (eds.), *Social Policy and Sociology*. New York: Academic Press.
- Sellin, T. and M. Wolfgang
1964 *The Measurement of Delinquency*. New York: Wiley.
- Tannenbaum, Frank
1938 *Crime and the Community*. New York: McGraw-Hill.
- U.S. Bureau of the Census
1970a "Victim recall pretest (Washington, D.C.): household survey of victims of crime." Mimeographed. Suitland, Md.
- 1970b "Household survey of victims of crime: second pretest (Baltimore, Md.)." Mimeographed. Suitland, Md.
- 1975a *Statistical Abstract of the United States*. Washington, D.C.: U.S. Government Printing Office.
- 1975b "Comparison of the second stage ratio factors produced in the National Crime Survey and the Current Population Survey." Mimeographed memorandum. Suitland, Md.

Williams, Jay and Martin Gold

1972 "From delinquent behavior to official delinquency." *Social Problems* 20:209-28.

Wolfgang, Marvin

1963 "Uniform crime reports: a critical appraisal." *University of Pennsylvania Law Review* 111:708-38.

Wolfgang, Marvin and Franco Ferracuti

1967 *The Subculture of Violence: Toward an Integrated Theory in Criminology*. London: Tavistock.

Wolfgang, Marvin, Robert Figlio and Thorsten Sellin

1972 *Delinquency in a Birth Cohort*. Chicago: University of Chicago Press.

COMMENTS

ON THE INTERACTION BETWEEN A MOVEMENT AND ITS ENVIRONMENT*

(COMMENT ON JENKINS AND
PERROW, ASR APRIL, 1977)

In arguing that "the political environment surrounding insurgent efforts alternatively contains them or makes them successful," the Jenkins-Perrow (1977: 249) paper ignores the reciprocity of the influence cycle between a movement and its environment. This is a particularly serious oversight in attempting to understand the history of the United Farm Workers' (UFW) struggle, and rather puzzling in a treatment claiming to be in the tradition of Oberschall and Tilly, both of whom emphasize what the former calls a "dynamic system analysis" (Oberschall, 1973: 26). The Jenkins-Perrow paper compares the mobilization attempts of the National Farm Labor Union (NFLU) and the UFW, finds them to have been essentially similar in terms of goals, tactics and obstacles, and then concludes that the critical difference between the former's failure and the latter's success was the nature of the political environments each encountered. The authors completely ignore the most recent five years of the UFW's struggle, but they end by suggesting that the post-1972 future looks dim for that movement and other challenging groups (p. 267). My criticisms here will focus on (1) the differences between the NFLU and the UFW, (2) the authors' environmental determinism, and (3) the predictive failure of their theory in the light of post-1972 history.

(1) *Dissimilarities between the NFLU and UFW.* Despite the numerous similarities between these farmworker mobilization attempts, two important differences must be stressed in the course of any serious comparison: ethnicity and the use of a national boycott. The former union was composed primarily of white, Southern, fundamentalist, Dust Bowl refugees whose churches were quite indifferent to worldly concerns, while the UFW is pre-

dominantly Mexican-Catholic in membership and culture with strong ties to the activist wing of the U.S. Catholic church. The ethnic and religious homogeneity of the UFW's membership has been deftly used by Chavez and his cadre to encourage critical support from the post-Vatican II church (Levy, 1975). Two former officials of the NFLU (Ernesto Galarza and Philip Vera Cruz) have emphasized the point that the earlier union never seriously attempted a national boycott (Grubbs, 1975). Part of Chavez's genius has been his recognition of the fact that his mobilization attempt had to be carried out on two fronts simultaneously, and he long ago named his two watchdogs "Huelga" (translation: "strike") and "Boycott" to symbolize this duality. The interchange between the fields and the cities has been well orchestrated by the UFW's leaders, with farmworkers moving to Detroit, New York, Boston and other major cities in order to contribute authenticity to the picket lines, and educated volunteers moving to rural areas to contribute their expertise to hiring hall administration.

A third difference between the two mobilization attempts is the UFW's ample supply of volunteer lawyers and law students, invaluable resources for checking grower and Teamster arrogance via lawsuits and free advice to farmworkers. Jerry Cohen, the UFW's chief attorney, assured me (in a taped interview) that Chavez's low key style of leadership has been critical in attracting and retaining this legal assistance.

(2) *Environmental determinism.* A second major flaw in the Jenkins-Perrow paper is its simplistic and monocausal treatment of the political environment as the determining factor distinguishing the NFLU's failure from the UFW's success. Why, one might ask, were the Indian, Women's and Gay Liberation movements, for example, less successful (assuming they were) when they emerged in the same political environment? Again, the authors are trapped by their own proclivity for monocausal explanations, and fail to emphasize the creative use of the favorable environment by the UFW's leaders. The paper suggests that the UFW was launched and automatically sustained by a liberal coalition (p. 250), when, in fact, that coalition was frequently directed by Chavez and his cadre. One need only mention Chavez's famous 25-day fast in 1968 which

* This paper emanates from research sponsored by the Ford Foundation which permitted the author to reside in the rural areas of California for the year 1974-75 while collecting data on the grower-Teamster-UFW conflict. An anonymous reviewer's comments improved the style of this paper.

many informed observers consider a turning point both for internal union cohesion and subsequent boycott success (Taylor, 1975; Levy, 1975; Walsh, 1977). The religious atmosphere of the fast alienated many labor, liberal and radical supporters, but it also attracted R. F. Kennedy, the national news media, and numerous rather conservative religious organizations to the side of this Mexican-American Ghandi. The chief attorney for the UFW, Jerry Cohen, summarized the effect of Chavez's fast on the organization itself:

Cesar was mad. There had been a lot of loose talk about violence. . . . They had to find out who had the balls, and he showed them. He scared hell out of them. . . . He just said the union was committed to non-violence, and then started fasting. The people responded like "God, what is this guy doing?" . . . What a fantastic cement that fast was . . . it was an amazing organizing tool. (Taylor, 1975:225)

The UFW, like any social movement, had to make strategic decisions about which potential resources were to be used, and how others were to be created. Some of those decisions were critical. Kushner (1975), for example, criticizes Chavez for refusing the help of the Communist Party (CP), but it is unlikely that organized religion or organized labor would have remained the loyal allies they proved to be had the UFW also accepted support from the CP. Indeed, the growers and local law enforcement officials broke into the UFW offices to discover evidence linking the UFW and the CP, but when nothing turned up, federal officials cooperated in a cover-up (Walsh, 1977). The alleged neutrality of the political environment vis-à-vis farmworker insurgency ended very quickly with Nixon's inauguration in 1968, but the Jenkins-Perrow paper does not even mention the Colson memos instructing the Justice and Labor Departments not to intervene in the Teamster-UFW dispute unless their actions would be harmful to the UFW, the assassination plot on Chavez covered up by the Treasury and Justice Departments, or other evidence contradicting the neutrality hypothesis in the 1968-72 period (cf. Levy, 1975; Walsh, 1977). My point is not to deny the importance of the political environment—especially for the launching of this insurgency attempt—but only to insist that other factors such as ideology and leadership were also crucial in determining its outcome.

(3) *Predictive failure.* The critical test of any hypothesis is prediction. The authors take us only to 1972, and warn that "even the gains of the past may be endangered" (p. 267) because the political environment was looking more bleak. Indeed, their prognosis was endorsed by

their favorite source with an obituary on the movement entitled "Is Chavez Beaten?" (*New York Times Magazine*, 9/15/74). In fact, however, the UFW has emerged from its struggle with the growers and their Teamster allies¹ as *the* farmworkers' union in the Southwest—and this during a period of U.S. history some might characterize as "reactionary." Observers generally agree that the Teamster-UFW pact of March 10, 1977 represented a decisive victory for the Chavistas (cf. *Wall Street Journal*, 3/11/77), and U.S. Labor Secretary Ray Marshall recently promised the UFW that his office would work to modify the National Labor Relations Act (NLRA) in accordance with the new union's demands (cf. *Fresno Bee*, 8/28/77). Such data, in addition to many earlier examples of the UFW's influence on organized labor and organized religion (cf. Kushner, 1975), present a serious challenge to Jenkins-Perrow's environmental determinism.

The critical turning point in the UFW's struggle for survival came with the election victory of Jerry Brown to succeed former California governor and outspoken critic of the Chavistas, Ronald Reagan, in November, 1974. Although the Jenkins-Perrow paper minimizes the importance of insurgents' effects on the political environment (e.g., p. 266), Brown was *deeply* influenced by Chavez's movement years before assuming the governorship and openly criticized as "a friend of Cesar Chavez" by his Republican opponent during the 1974 gubernatorial campaign.² The new governor made it his first order of business to push through the California Legislature the now famous Agricultural Labor Relations Act (ALRA) within the first year of his incumbency—assuring farmworkers the right to vote in collective bargaining elections—and thus shifted the power of decision to the field workers themselves. At this new stage in the conflict, support from the movement's social base was crucial, and Chavez made his popular

¹ The authors' seem to justify their exclusion of the Teamsters by saying that this competing union entered in 1973 (cf. p. 254), but the Teamsters entered the conflict as early as 1966, and again in the summer of 1970, at the invitation of the growers. The seriousness of their 1970 involvement is well documented in the *Englund v. Chavez* decision of the California Supreme Court (12/29/72).

² The relationship between the California governor and Chavez is evidenced more recently in Brown's selection of Leroy Chatfield (a former UFW official) as his presidential campaign chairman, his use of UFW boycott offices for campaign centers in various states, and Chavez's introduction of the Brown nomination to the Democratic Convention in the summer of 1976.

"1,000-mile walk" through the rural areas of California in the 90-day interim period between passage of the ALRA and its becoming law. The purpose of this difficult and symbolic walk was to inform farmworkers of their voting rights in the face of grower-Teamster intimidation tactics. Contrary to the Jenkins-Perrow paper's global dismissal of internal movement considerations (e.g., p. 251), Chavez's efforts seem to have been influential in obtaining election victories at hundreds of ranches, and the elimination of the Teamsters from the fields.

In conclusion, I am in basic agreement with the authors' argument that environmental factors have been too much neglected in standard social movement literature. My objection is that Jenkins and Perrow go to the other extreme in neglecting important intramovement characteristics and, especially, the *interaction* between a movement and its environment. The course of this particular movement's history would have been very different without Chavez and Jerry Brown in the leadership positions they occupied, and if such personal attributes are essentially non-sociological in character, one can only echo Lenski's (1966:438) admonition that "sociological theories must find a place for them." It is hardly an improvement in sociological eyesight to counter traditional myopia (focusing exclusively on internal movement characteristics) with an equally deficient hyperopia (neglecting such details for the "big picture"). However conceptually inconvenient the dynamic reciprocity between and among social structures, deprived groups, and "great persons" may be, it is still more fruitful to attempt to construct a process model including them all than to settle for more simple and misleading monocausal explanations such as the one offered us in the Jenkins-Perrow paper.

Edward J. Walsh
Pennsylvania State University

REFERENCES

- Englund v. Chavez
1972 Sup., 105 (Cal. Rptr. 521).
Fresno Bee
1977 August 28.
Grubbs, Donald H.
1975 "Prelude to Chavez: the National Farm Labor Union in California." *Labor History* 16: 453-69.
Kushner, Sam
1975 *Long Road to Delano*. New York: International Publishers.
Lenski, Gerhard
1966 *Power and Privilege*. New York: McGraw-Hill.
Levy, Jacques
1975 Cesar Chavez: Autobiography of La Causa. New York: Norton.
New York Times Magazine
1974 September 15.
Oberschall, Anthony
1973 *Social Conflict and Social Movements*. Englewood Cliffs: Prentice-Hall.
Taylor, Ronald
1975 *Chavez and the Farm Workers*. Boston: Beacon.
Wall Street Journal
1977 March 11.
Walsh, Edward
1978 "Mobilization theory vis-à-vis a mobilization process: the case of the United Farm Workers' movement." In Louis Kriesberg (ed.), *Research in Social Movements, Conflicts and Change*. Greenwich, Ct.: Ai Jai Press. In press.

THE ECOLOGICAL APPROACH AND COMMUNITY LEADERSHIP

(COMMENT ON GRIMES, BONJEAN,
LYON AND LINEBERRY,
ASR AUGUST, 1976)

The study by Grimes et al. (1976) has a worthwhile aim: to obtain measures on different facets of community leadership for a sufficiently large sample of communities to permit multivariate analyses of the conditions producing variation in leadership arrangements. There are, however, certain conceptual and methodological problems in their paper which require comment. Our objections address three issues: (1) their claim to evaluate an ecological approach to community power studies and the criteria they use to make this evaluation; (2) their assessment of multidimensionality in local leadership arrangements; and (3) their derivation of independent variables and thus hypotheses from a factor analysis of U.S. counties.

An Ecological Approach?

We think it first important to point out that the only sense in which Grimes et al. assess an ecological approach to community power is that derived from Robinson (1950) which pertains to the use of quantitative areal data in comparative multivariate studies of communities. This usage of "ecological" bears little relation to ecological theory (see Hawley, 1963; and Lincoln, 1976 for applications of such theory to problems of community power). Consequently, the analyses they perform fail to speak to any such theoretical issues. These analyses and the authors' inferences from

them, moreover, are flawed in ways which make it difficult to discern exactly what it is the authors have learned from their examination of these data. We outline some of these problems below.

The authors combine various measures of community composition in regression equations predicting three dimensions of community leadership and express disappointment in the R^2 's—although these, which range up to .43, are not trifling by prevailing standards. Such small proportions of variation accounted for suggest to them that "the explanatory power of the ecological approach" is low because it "necessitates the measurement of a qualitative phenomenon (whether it be 'power structure,' 'party competition,' 'democratization,' or similar phenomena) with entirely quantitative indices" (p. 722). We are perplexed by this statement. Are they suggesting that only an exploratory case study strategy can reveal the inner workings of community power structures? Or, as their examples suggest, do they mean that only nominal distinctions can be drawn among types of political arrangements, ruling out higher (i.e., quantitative) levels of measurement? Our first reaction is that (amount of?) party competition and (degree of?) democratization are hardly illustrations that support their point. But grant them this claim for the sake of argument; why not then use a qualitative breakdown of their leadership measures and adopt a statistical procedure which can cope with qualitative dependent variables? The very similar comparative analyses of local power structure by Aiken (1970) and Walton (1967) do just this.

The above is just one example of a tendency of these writers to identify a minor methodological obstacle as a flaw intrinsic to the ecological approach. A similar example is their complaint that past comparative studies may have had low predictive power because of a failure to allow for nonlinear effects in models relating community traits to power structure indicators (pp. 721, 723). Surely, given the ease with which this hypothesis might be evaluated in the course of their own analysis, the next step will be for them to test for the presence of such effects themselves? But nothing of the sort is done here. Nor, in the face of their own neglect of precisely the same issues, do the authors seem aware of the contradiction in their critique of other work.

As these points imply, there is an exaggerated concern in this study with the magnitude of R^2 which is the standard against which the ecological approach is evaluated and found wanting. At another juncture they note: "The adjusted R^2 also provides a very rough assess-

ment of Clark's general hypothesis that differentiation in community structure is associated with a more pluralistic leadership structure" (p. 717). It is debatable as to whether the measures they use here can be construed as indicators of differentiation, but the main question we have is how the sheer ability to predict pluralism from community composition says anything about Clark's hypothesis. As the statement stands, one must presume that even negative effects on pluralism of the differentiation measures would confirm Clark's hypothesis given a sufficiently high R^2 .

The Issue of Multidimensionality in Leadership Structure

Following earlier work by Bonjean (1971) and others, the authors identify four dimensions of local leadership structure. *Legitimacy* and *visibility* are measured for 17 communities studied by the authors and others; *scope of influence* is equated with Clark's measure of decentralization; and a last dimension, *consensus*, is mentioned, then inexplicably dropped from further consideration. Grimes et al. (1976: 713) suggest that, under the hypothesis of unidimensionality, pluralist communities should have visible, legitimate, and decentralized leadership; the opposite pattern representing elite dominated communities. We, however, do not find so obvious the easy equation of democracy with pluralism on which this typology depends. Visibility and legitimacy are not *prima facie* expressions of the degree to which diverse interest groups compete in the political arena. If most observers imply that pluralist systems are more democratic, they do so out of consideration of certain causal mechanisms (e.g., Dahl's [1961: 163] "indirect influence") whereby elite competition gives rise to rank-and-file power. But we suspect that few would defend the premise of the Grimes paper that the occupancy of formal, public posts—conceded to be an element of formal democracy—is thereby an expression of pluralism as well.

To assess the hypothesis of unidimensionality, Grimes et al. (1976) regress these three leadership measures on a set of seven community traits. We have serious doubts that the equation for decentralization should be compared with the other two since it is based on an altogether different sample: the NORC Permanent Community Sample of 51 cities. The Grimes et al. (1976) sample of 17 communities greatly overrepresents the South, a fact having significant implications for the leadership arrangements one would expect them to have (see Aiken, 1970). But as regards legitimacy

and visibility, we are again perplexed that the authors never present that datum which bears most directly on the matter of whether these tap the same underlying construct: the correlation between them. They themselves assert that this hypothesis rests on the "degree of collinearity" (p. 720) among the leadership dimensions but still conceal from us that piece of information which speaks explicitly to the question. The most rigorous approach here would have been to treat leadership structure as an unobserved construct in a structural equation model specifying community traits as determinants and the specific leadership dimensions as multiple indicators. Lincoln (1976) takes this approach in addressing related questions pertaining to the dimensionality of community policy outputs.

Factor Analysis and Hypothesis Formulation

Finally, we must address the hypotheses which the authors claim are derived from the literature. On the contrary, what they have done is to identify empirically properties on which communities differ by means of a previous factor analysis of 79 variables for all U. S. counties. They have then searched the literature to find some precedent for looking at the association between each of these characteristics and the leadership indicators. Thus, the seven hypotheses they derive, which link pluralism of leadership structure to community social and economic composition, represent an ad hoc justification of predictors selected by purely empirical procedures. The conventional rationale for this kind of strategy is that the factors represent basic patterns of variation in community structure; hence, their association with leadership arrangements is worthy of scrutiny. But aside from the old problem that underlying dimensions are only as good as one's set of initial measurements, two additional problems are present. First, the factor analysis was performed not on a set of community units, but on the universe of U. S. counties. Perhaps it thus tells us something about territorial differentiation in toto. Given the large expanses of rural and unpopulated territory included, it is quite uncertain as to whether it says anything meaningful about differences among communities. Second, measures of local revenues and expenditures appeared in the correlation matrix from which the factors were derived. Since these are generally viewed as outputs of community political systems, Grimes et al. (1976) elect to exclude from

their list of compositional variables a factor corresponding to this set of measures. What they fail to recognize is that the entire set of factors is simultaneously determined from the correlations in the original matrix. Thus the factors they select as predictors are defined to some degree by the presence of the budgetary data.

Thus, despite interesting data, we find this study flawed on both conceptual and methodological grounds. As a final note, we would hope that we have drawn attention to the ambiguities involved in the characterization of aggregate data as "ecological" so that genuinely ecological approaches to the study of community structure may be readily discerned.

James R. Lincoln

Jon Olson

University of Southern California

REFERENCES

- Aiken, Michael
1970 "The distribution of community power: structural bases and social consequences." Pp. 487-525 in Michael Aiken and Paul E. Mott (eds.), *The Structure of Community Power*. New York: Random House.
- Bonjean, Charles M.
1971 "Dimensions of power structure: some problems in conceptualization and measurement." Pp. 19-42 in F. Wirt (ed.), *Future Directions in Community Power Research*. Berkeley: University of California Institute of Governmental Studies.
- Dahl, Robert A.
1961 *Who Governs? Democracy and Power in an American City*. New Haven: Yale University Press.
- Grimes, M. D., C. M. Bonjean, J. L. Lyon, and R. L. Lineberry
1976 "Community structure and leadership arrangements: a multidimensional analysis." *American Sociological Review* 41:706-25.
- Hawley, Amos H.
1963 "Community power and urban renewal success." *American Journal of Sociology* 68:422-31.
- Lincoln, James R.
1976 "Power and mobilization in the urban community: reconsidering the ecological approach." *American Sociological Review* 41:1-15.
- Robinson, W. S.
1950 "Ecological correlations and the behavior of individuals." *American Sociological Review* 15:351-7.
- Walton, John
1967 "The vertical axis of community organization and the structure of power." *Social Science Quarterly* 48:353-68.

ON INEQUALITY IN THE MILITARY*

(COMMENT ON BUTLER, ASR
OCTOBER, 1976)

Butler (1976) analyzed racial differences in the mean number of months from enlistment to promotion to current rank for all enlisted men at or above the rank of corporal in the U.S. Army in 1973. The paper includes several tables in which AFQT (Armed Forces Qualification Test) scores, years of schooling, and military occupation (technical or nontechnical) were first separately and then simultaneously controlled among black and white enlisted men at three ranks or combinations of ranks. For example, Table 1 (Butler's Table 2) shows mean months to promotion to current rank by race and AFQT score; E4 is the lowest rank and E9 is the highest. By inspection it is evident that within the same broad ability class and at each rank black men took somewhat longer than white men to achieve their current rank. Moreover, except in the case of rank E4 among white men, it took rank incumbents with high AFQT scores longer to enter their current rank than it took men with low AFQT scores. Results similar to these were obtained by Butler when all of the controls were introduced simultaneously. With reference to these tabulations, Butler (p. 807) writes:

Taking promotion time as a measure of inequality, this paper presents data which suggest that universalistic criteria are not sufficient to explain black inequality vis-à-vis promotion time. When black and white enlisted men are matched on key uni-

versalistic criteria (civilian education, Armed Forces Qualification Test, and occupation type), blacks consistently take more time to make grade than whites. The findings suggest the salience of the notion of ascription: when ascription is based on race, it becomes a component part of racism. Thus, one unavoidably is led by the data in this paper to conclude that the black enlisted man is subject to inequality, which is not the result of failure to meet universalistic criteria but, rather, the result of the racist actions of real-life people.

Later he writes:

Common sense would lead one to expect that personnel in the high AFQT category would move faster than those in the low AFQT category, but this expectation was not realized. One interpretation is that people who scored in the high mental category are less likely to conform to Army expectations and norms. If this line of conjecture is accurate, it may also be that within mental and educational levels, blacks compared to whites are less susceptible to subservient (or from the establishment's viewpoint accommodative) behavior toward superiors. (p. 816)

These two excerpts do not exhaust Butler's interpretations of his data, but I believe they are representative.

Butler deserves substantial credit for his imaginative application of administrative data to a significant research and policy issue. Unfortunately, differentials in mean months to promotion bear no necessary relationship to Butler's conclusions and for this reason can neither confirm nor disconfirm them. The problem is that there is no closure in the military population, either with respect to movement out of the military into the civilian population (or death) or with respect to movement between ranks. There is no self-evident interpretation of the temporal differences between black and white rank incumbents in months in service to make current rank because persons occupying a given rank do not represent the experience of all persons who have been promoted to that rank, nor of all persons exposed

* This was prepared while the author received support from the National Science Foundation (Grants No. GI-31604 and GI-4436) and the National Institute of Mental Health (Grant No. MH-06275). I thank Morris Janowitz, William H. Sewell, Franklin Wilson, Halliman H. Winsborough, and Wendy Wolf for their comments on earlier drafts. The opinions expressed herein are those of the author.

Table 1. Mean (\bar{X}) Months in Service to Make Grades E4, E5-E6, E7-E9, Controlling for AFQT Mental Groups by Race*

Grade(s)	Mean (\bar{X}) Months in Service			
	Black		White	
	AFQT Low	AFQT High	AFQT Low	AFQT High
E4	18 (17,016)	23 (1,657)	16 (63,581)	14 (44,368)
E5-E6	60 (22,950)	93 (7,096)	57 (60,048)	62 (52,164)
E7-E9	165 (5,841)	181 (5,311)	161 (16,731)	173 (30,448)

* N in parentheses; months rounded to nearest whole.

to the "risk" of promotion. Some persons who were promoted to the given rank will subsequently have left the military, while others will have been promoted to higher ranks. Moreover, months in service to make current rank is not a valid inverse measure of the likelihood of promotion to that rank because it does not take account of the experience of persons who might have been promoted to that rank but were not.

For the usual logic of multivariate analysis to apply, one has to believe that persons have (more or less) equal exposure to the stratification regime; that is, that the population is closed, or that leaving the population does not occur differentially with respect to the phenomenon under investigation. (This is a significant problem in comparing the social and economic standing of blacks and whites or of men and women in the civilian labor force.) The Army is not a closed population, and the problem is further compounded by the fact that movement upward or downward in the ranks probably occurs differentially (Nordlie, Thomas and Sevilla, 1975). In short, the mean time to promotion to current rank of current rank incumbents probably does not accurately reflect Army promotional policies. Regardless of statistical controls, Butler's data do not reflect racial differentials in promotion conditional on equivalent exposure to the risk of promotion.

Suppose that black men with given qualifications have about the same chance of promotion as equally qualified white men, while time to last promotion is negatively correlated with the likelihood of further promotion. Those suppositions suggest that the *current* incumbents of a given rank would include a disproportionate number of those men who took longest to achieve it. That alone might account for the negative relationship between AFQT scores and mean time to promotion to current rank. That is, the higher scoring incumbents of a given rank would include a relatively larger number of men who had been passed over for promotion than the lower scoring incumbents of the same rank because the most promising of the higher scoring incumbents would already have been promoted. Moreover, consider the fact that whites have better opportunities than blacks for socioeconomic rewards in civilian employment (Featherman and Hauser, 1976). Given these differential opportunities we might expect relatively more of the less meritorious whites than of the less meritorious blacks to leave the Army, and this differential in retention could account for the racial difference in mean time to promotion among current rank incumbents. Finally, this argument suggests

that the distribution of rank by race might be more favorable to blacks than is indicated by the array of mean times to promotion.

These are not the only factors which might invalidate Butler's interpretation of the Army data, but I suspect they are among the more important. There are others worth mentioning. What would the data look like if the Army were in the process of moving from discriminatory to nondiscriminatory promotion policies? In what ways does a military force which was recruited by a mixture of draftees and career-enlisted men complicate the picture? How are differentials in the size and composition of entering cohorts of enlisted men and of their chances for promotion affected by the differing conditions of the pre-Vietnam and Vietnam war eras? Does a control for military occupation adequately reflect differentials in promotional opportunities between branches of service? All of these phenomena may be reflected in Butler's Army data, but I shall not pursue them further here.

Butler's data may be rearranged to support the main argument. From the numerical counts in Tables 5a, 5b, 6a, 6b, 7a, and 7b, I have constructed Table 2, which shows the percentage of men in each rank by race, AFQT, years of schooling and military occupation.¹ Within levels of schooling, AFQT, and military occupation, black enlisted men without exception are *less* likely than whites to be in rank E4 relative to E5 through E9. The differences between the races are especially marked among men with high AFQT scores. With only one exception (low AFQT, nontechnical occupation, middle education), black men were *more* likely than whites to be in grades E7 to E9 relative to the numbers in E4 through E6. Again, the racial differences in rank are largest among the men with high AFQT, where the (opposite) racial differences in mean time to promotion were greatest.

Other findings also appear to contradict Butler's interpretation of his data. Most noteworthy is the consistent finding that the high AFQT men are less likely to be in rank E4 and more likely to be in ranks E7 through E9 than men of low AFQT scores with the same race, education and military occupation. The lack of substantial differences in rank between men of middle and high educational attainment

¹ I have assumed that the figure of 11,548 given in Table 7b, row 1, column 2, should be 1,548. That is consistent with other tables. I have not made any other changes, but there are other smaller inconsistencies between tables. Professor Butler reports that these inconsistencies were contained in the original tabulations supplied by the Army.

Table 2. Military Rank by Race, AFQT Score, Education, and Military Occupation*

AFQT	Occupation	Rank	Black			White		
			Low Education	Middle Education	High Education	Low Education	Middle Education	High Education
LOW	Nontechnical	E4	40.1%	37.6%	41.5%	50.4%	44.7%	54.8%
		E5-6	54.3	48.4	43.7	45.5	38.9	34.5
		E7-9	5.7	14.1	14.8	4.2	16.4	10.7
			100.0	100.0	100.0	100.0	100.0	100.0
			(8,575)	(20,517)	(2,737)	(34,489)	(54,994)	(6,500)
	Technical	E4	38.0	31.8	36.5	47.3	37.4	43.3
		E5-6	54.3	51.4	47.9	47.4	46.7	44.6
		E7-9	7.7	16.8	15.6	5.3	16.0	12.1
			100.0	100.0	100.0	100.0	100.0	100.0
			(2,935)	(9,231)	(1,866)	(12,194)	(27,675)	(4,508)
HIGH	Nontechnical	E4	11.0	9.9	20.1	37.3	32.4	49.7
		E5-6	66.8	48.6	36.6	47.0	38.3	32.4
		E7-9	22.1	41.5	43.3	15.7	29.3	17.9
			100.0	100.0	100.0	100.0	100.0	100.0
			(1,883)	(6,041)	(962)	(11,802)	(44,397)	(13,511)
	Technical	E4	11.6	9.5	20.7	32.6	29.1	42.1
		E5-6	61.4	50.0	39.3	49.3	44.0	41.9
		E7-9	27.0	40.5	40.0	18.2	26.9	16.0
			100.0	100.0	100.0	100.0	100.0	100.0
			(760)	(3,457)	(907)	(6,517)	(35,388)	(15,365)

* From Butler, 1976; suitable counts are in parenthesis.

suggests the need to control for mode of entry into the military service. Most importantly, conditional on promotion to rank E4, Butler's data show that blacks are more likely than whites of similar AFQT, schooling and military occupation to be in the higher enlisted ranks. Also, the differentials favoring blacks in the chance of occupying a high rank are larger where the apparently opposite tendencies in mean time to promotion are greater.

Do these data suggest a sociological interpretation of Army promotional practices in the enlisted ranks which is opposite from that made by Butler? That is, are blacks favored over whites by Army promotional practices? (Note that many of the racial differentials favoring blacks in Table 2 are quite large.) The answer, I think, is *no*, and for the same reason that I cannot accept Butler's interpretation of his data. Neither the data on rank by race, AFQT, schooling, and occupation or the data on mean months to promotion by rank, race, AFQT, schooling, and occupation are adequate for a sociological diagnosis of Army promotional practices. These data are a cross-sectional residue of fairly long-term processes of recruitment, promotion, and retention which cannot accurately reflect the operation of any of those processes at any time. The seeming opposite tendencies observable in the rank distribution and in mean time to promotion both reflect the inadequacy of the research design to answer the main question of sociological interest—the degree of racial inequality in Army promotional practices.

Robert M. Hauser
*University of Wisconsin,
Madison*

REFERENCES

- Butler, John Sibley
1976 "Inequality in the military: an examination of promotion time for black and white enlisted men." *American Sociological Review* 41:807-18.
- Featherman, David L. and Robert M. Hauser
1976 "Changes in the socioeconomic stratification of the races, 1962-1973." *American Journal of Sociology* 82:621-51.
- Nordlie, Peter G., James A. Thomas, and Exequiel R. Sevilla
1975 *Measuring Changes in Institutional Racial Discrimination in the Army*. Technical Paper 270. Arlington, Va.: U.S. Army Research Institute for the Behavioral Sciences.

COMPARING MODELS OF MOBILITY*

(COMMENT ON McCANN,
ASR FEBRUARY, 1977)

McCann (1977) offers a model of mobility in terms of which he can describe and compare aspects of British and Danish mobility. He notes that differences exist between his model and my own (Levine, 1972), but dispatches the one model (and elaborates the other) without due attention to either technical or substantive criteria by which competing models (applied to the same data) are normally compared. McCann's argument illustrates a general weakness in the use of mathematical models as tools for the description and assessment of mobility, viz., that the development and elaboration of new models has been pursued without due test and examination of the validity of their premises.

Attending first to the technical contrast, McCann (1977) notes that my earlier model obtains a better fit to Glass and Hall's (1954) British data, in particular by reducing residuals with respect to predictions of diagonal values (predictions of immobile men). He dismisses this by noting that the model "accords special weight to the diagonal values." This statement of contrast is misleading in so far as it suggests the model sacrifices parsimony for fit, piling-on parameters (weights) until the residuals go away. Quite the opposite is the case. In fact, the McCann model itself comes close to "curve fitting" in the specific sense that, of the 16 degrees of freedom available for interaction, the model spends 14. Two more and the fit would have to be perfect as it takes only 16 parameters to fit a table with 16 degrees of freedom. By contrast, the model which McCann rejects spends only 5 degrees of freedom and, nevertheless, obtains better fits (to both the British and Danish data). Table 1 shows both sets of predictions for the British data (my own being revised to match McCann's aggregation of the original categories). Table 2 shows the corresponding predictions (my own previously unavailable) for the Danish data (Svalastoga, 1959). Reporting standardized fit as chi-square divided by its degree of freedom, the earlier model obtains $13.5/11 = 1.2$ compared to McCann's $41.8/2 = 20.9$ for the British data and $23.5/11 = 2.1$ compared to McCann's $31.8/2 = 15.9$. Chi-squares, adjusted chi-squares, and degrees of freedom

* I am indebted to David K. Miller, Robert Z. Norman, and David Luchini for comments and editorial suggestions.

Table 1. Observed British Mobility Frequencies^a

Origin Status	Destination Status British Sons				
	1	2	3	4	5
1	50 (48.1) [50.3]	45 (42.7) [42.6]	8 (14.3) [13.6]	18 (18.9) [18.8]	8 (5.0) [5.5]
2	28 (36.6) [29.3]	174 (149.5) [173.9]	84 (90.6) [81.9]	154 (166.8) [148.5]	55 (51.5) [62.0]
3	11 (7.2) [9.9]	78 (87.9) [87.1]	110 (86.7) [101.1]	223 (232.8) [213.3]	96 (103.4) [108.4]
4	14 (6.8) [12.4]	150 (148.3) [142.5]	185 (195.2) [192.5]	714 (709.0) [734.7]	447 (450.7) [429.0]
5	0 (4.5) [2.7]	42 (60.4) [43.6]	72 (72.8) [71.7]	320 (301.0) [314.3]	411 (406.3) [414.2]

^a McCann model predicted frequencies (in parentheses); Levine model predicted frequencies (in square brackets).

all favor the earlier model. By criteria of both fit and parsimony, McCann's model is wanting.

Only one parameter of the model may be said to have special weight on the diagonal. While McCann prefers not to so favor the diagonal, nonetheless the choice of one or another model (and thereby of one or another theory of the principal characteristics of mobility) is an issue to be decided by test, not by taste. He must offer evidence and, at least in principle, predictions based on the premises he prefers.

While McCann's model is in these strict senses inferior, the technical distinctions should not be overdrawn. The real technical distinction between McCann's model and my own is, I suspect, quite minor and lies in the shape of functions we chose as embodiments of our particular theories. McCann uses Beta distributions; I use a log-quadratic function of the distance between social classes. McCann's choice has relatively rounded peaks (continuous in its derivatives); my own has a cusp, i. e., a very sharp peak. This distinction between the

Table 2. Observed Danish Mobility Frequencies^a

Origin Status	Destination Status Danish Sons				
	1	2	3	4	5
1	18 (18.4) [16.4]	17 (16.4) [17.4]	16 (14.4) [15.7]	4 (6.6) [6.5]	2 (1.2) [1.6]
2	24 (34.5) [31.9]	105 (80.8) [96.6]	109 (116.4) [111.4]	59 (72.2) [60.7]	21 (14.2) [18.9]
3	23 (18.8) [23.8]	84 (103.1) [92.4]	289 (258.7) [282.7]	217 (234.9) [198.9]	59 (56.5) [77.7]
4	8 (5.8) [8.7]	49 (47.5) [44.3]	175 (191.8) [175.1]	348 (327.4) [367.8]	198 (205.4) [183.5]
5	6 (1.5) [2.3]	8 (14.5) [14.8]	69 (75.0) [73.4]	201 (188.3) [197.0]	246 (250.7) [247.2]

^a McCann model predicted frequencies (in parentheses); Levine model predicted frequencies (in square brackets).

models seems minor in so far as, by my reading of McCann's argument, his argument is not particularly wedded to the Beta distribution. Moreover, I would conjecture that with only a moderate display of mathematical virtuosity, I could rewrite my equations to match the specifications—and thus appear to support the premises of—McCann's verbal argument while he in turn could rewrite his equations to match the verbal specifications of my own.

Therein lies the point which is more important than the strictly technical distinctions between models. There are theoretical questions at issue when models are compared and the issues should not depend solely on undefended preferences (McCann's criterion) nor on nonessential technical distinctions. The theoretical contrast between McCann's argument and my own is, as McCann notes, that I formulate mobility as a matching process, man to job and job to man, governed by the distance between the status of origin and the status of arrival, while McCann formulates mobility as a process of occupational attainment governed primarily by the qualities of the candidates. In my formulation the two parties to the man-job arrangement, i.e., the man seeking the position and the position (or controllers of the position) seeking the man, are treated as dual entities. This makes position-person matching similar in this regard to a market-bartering phenomenon, which matches the two parties to a transaction. There are at least three substantively distinct alternatives. Is mobility an attainment process in which men are the actors qualifying themselves for and competing for jobs (McCann)? Is it a recruitment process in which the jobs themselves are "actors" shaping, selecting, and competing for the available men (White's [1970] *intra-organizational* model)? Or is it a dual process in which both sides of the "incumbency" (White's term) are matched?

These are differences of substance (see White, 1970: Chap. 1) which should not be decided a priori. Nor should they be decided by whether McCann (or I) did or did not allow a cusp to appear in our equations—unless that choice was firmly dictated by our respective theories. The assumptions of the model must be evaluated and its accuracy vis-à-vis alternative models must be tested. Otherwise it is frivolous to elaborate the extensive implications of a model (as they are elaborated for Britain and Sweden and for structural and exchange mobility by McCann, 1977: 84–9).

The priority of test and comparison between the alternative assumptions and predictive power of competing models extends beyond the two-way contrast between McCann's model and my own. Are statuses to be treated

as interval-scale categories (both models) or only as ordered categories (consistent with Goodman, 1969: described by McCann, 1977)? This is a difference of substance. Which is correct; which is more parsimonious? Are the marginals of the classes of destination "logically prior to the allocation of sons" (McCann or Levine model), or are the marginals of the classes of destination determined by the allocation process itself (implied by Markovian models; e.g., Mayer, 1968; 1972; McFarland, 1970; McGinnis, 1968; Blumen et al., 1955; Prajs, 1955)?

Are the causes of mobility substantively different from the causes of *immobility* (as implied by the modified Markovian models) or is immobility simply one point on a continuum among moves of greater distance (consistent with McCann's [1977], McClendon's [1977], or Levine's [1972] methods)? Is differential mobility more a consequence of positive recruitment of favored persons or a consequence of discriminatory bars against others (White, 1963)?

These are substantive issues by which models can be classified. These are issues which can be decided or at least explored by testing the models. Which premises survive or prove to be stronger when the models (and thus, indirectly, the premises) are tested against each other? It would seem that before we practice the familiar litany of most mathematical modelers of mobility (. . . pick a model, then use it to analyze the Glass and Hall [1954] data and possibly Svalastoga's [1959], make a few comparative statements inferred from the model, and then comment on structural and exchange mobility under the assumptions of the model . . .)—before we practice this litany, there is a prior obligation to be met. We must be reasonably certain the model offered is not just "another model" or a more clever model. It must have clear links to theory and must survive tests against alternative models.

Joel H. Levine
Dartmouth College

REFERENCES

- Blumen, I., M. Kogen, and P. J. McCarthy
1955 The Industrial Mobility of Labor as a Probability Process. Cornell Studies of Industrial and Labor Relations, Vol. 6. Geneva, N.Y.: Humphrey.
- Glass, D. V. and J. R. Hall
1954 "Social mobility in Britain: a study of intergenerational changes in status." Pp. 177–217 in D. V. Glass (ed.), *Social Mobility in Britain*. Glencoe: Free Press.

- Goodman, Leo A.
1969 "How to ransack social mobility tables and other kinds of cross-classification tables." *American Journal of Sociology* 75:1-40.
- Levine, Joel H.
1972 "A two-parameter model of interaction in father-son status mobility." *Behavioral Science* 17:455-65.
- Mayer, Thomas F.
1968 "Age and mobility: two approaches to the problem of nonstationarity." *Annual Meeting of the American Sociological Association*, Boston.
1972 "Continuous time models of intergenerational mobility." Pp. 308-57 in J. Berger, M. Zelditch, and B. Anderson (eds.), *Sociological Theories in Progress*, Vol. 2. Boston: Houghton-Mifflin.
- McCann, J. C.
1977 "A theoretical model of the interpretation of tables of social mobility." *American Sociological Review* 42:74-91.
- McClendon, McKee J.
1977 "Structural and exchange components of vertical mobility." *American Sociological Review* 42:56-73.
- McFarland, David D.
1970 "Intragenerational social mobility as a Markov process: including a time-stationary Markovian model that explains observed declines in mobility rates over time." *American Sociological Review* 35:463-76.
- McGinnis, Robert
1968 "A stochastic model of social mobility." *American Sociological Review* 33:712-22.
- Prais, S. F.
1955 "Measuring social mobility." *Journal of the Royal Statistical Society, Ser. A*: 55-66.
- Svalastoga, Kaare
1959 *Prestige, Class, and Mobility*. London: Heinemann.
- White, Harrison C.
1963 "Cause and effect in occupational mobility tables." *Behavioral Science* 8:14-27.
1970 *Chains of Opportunity: System Models of Mobility in Organizations*. Cambridge, Ma.: Harvard University Press.

NEW FROM COLUMBIA

FROM RITUAL TO RECORD

THE NATURE OF MODERN SPORTS

Allen Guttman

The evolution of modern sport and its special role in the modern world are detailed in this unique study. Using methods that range from literary criticism to statistical analysis, but concentrating on historical and sociological description, Guttman analyzes the trends toward specialization, equality, and record-seeking which distinguish modern sports from their predecessors. \$12.95

SOVIET CRIMINOLOGISTS AND CRIMINAL POLICY

Peter H. Solomon, Jr.

In this pathbreaking examination of the role of Soviet criminologists in the formation of Soviet criminal policy, the author challenges the conventional rubber-stamp image of specialist advisers to the Soviet government and suggests the need for a new interpretation of the scope of political process in the USSR. *Studies of the Russian Institute, Columbia University* \$20.00

To order send check or money order to Dept. JN. Individuals must enclose payment. Institutions may request billing.



COLUMBIA UNIVERSITY PRESS

Address for orders: 136 South Broadway, Irvington, New York 10533

ITEMS (Continued)

lowing new associate editors: ROBERT COLE, BEVERLY DUNCAN, ROSABETH KANTER, MURRAY MELBIN, ANTHONY OBERSCHALL, SAMUEL PRESTON, and DAVID SNYDER.

■ MURRAY MELBIN (Night as Frontier) is Professor of Sociology at Boston University. He is continuing research on the city as a system in 24-hour time. His study includes analysis of cycles and rhythms of behavior as well as interdependences among urban activities, and time-relations among cities. In 1972 he published *Alone and with Others: A Grammar of Interpersonal Behavior* (Harper and Row). Currently he is working on a manuscript entitled *Night as Frontier*, which will give the complete theoretical argument and evidence upon which the article appearing in the present issue of ASR is based.

■ BRIAN L. PITCHER (The Diffusion of Collective Violence) is a doctoral candidate in the Department of Sociology at the University of Arizona. His research interests include modeling trends in birth cohort fertility and models of international research. ROBERT L. HAMBLIN is Professor of Sociology at the University of Arizona. He is doing research on analytical models of political violence. His most recent publication (with John H. Kunkel) is *Behavioral Theory in Sociology* (Transaction Books, 1977). His earlier publications include *A Mathematical Theory of Social Change* (Wiley, 1973). JERRY L. L. MILLER is Associate Professor of Sociology at the University of Arizona. His research interests include epidemics of collective violence in Latin America.

■ RACHEL A. ROSENFELD (Women's Intergenerational Occupational Mobility) is Assistant Professor of Sociology at McGill University. Currently she is doing research on occupational careers, and is especially interested in the relationship between employment patterns and occupational achievements. She is beginning a study of labor market entry and early occupational achievements of a cohort of young men and women. Additionally, she is working (with François Neilsen) on substantive interpretations of differential equation models of social processes.

■ KARL L. ALEXANDER (Curriculum Tracking and Educational Stratification) is Assistant Professor at the Department of Social Relations, The Johns Hopkins University. He is engaged in studies of family contingencies on educational attainment, female labor force participation and socioeconomic achievements, IQ and the social organization of education, and postsecondary institutional differentiation. MARTHA A. COOK is Associate Research Scientist at the Center for Social Organization of Schools, The Johns Hopkins University. EDWARD L. MCDILL is Professor in the Department of Social Relations, The Johns Hopkins University. He is doing research on violence in schools, and cognitive and noncognitive outcomes of schooling. He edited (with James McPartland) *Violence in Schools: Perspectives, Programs, and Positions* (D. C. Heath, 1977).

■ RICHARD P. APPELBAUM (Marx's Theory of the Falling Rate of Profit) is Assistant Professor of Sociology at the University of California, Santa Barbara. In 1971 he published *Theories of Social Change*

(Rand McNally). His most recent work (with Jennifer Bigelow, Henry Kramer, Harvey Molotch, and Paul Relis) is entitled *The Effects of Urban Growth: A Population Impact Analysis* (Praeger, 1976).

■ ALBERT A. SIMKUS (Residential Segregation by Occupation and Race) is a Ph.D. candidate in the Department of Sociology and the Center for Demography and Ecology at the University of Wisconsin, Madison. He is engaged in research on the residential and social segregation of socioeconomic groups. He also is associated with the project "Occupational Changes in a Generation, II." He is studying also occupational mobility and assortative marriage in the U.S.

■ MICHAEL J. HINDELANG (Race and Involvement in Common Law Personal Crimes) is Professor in the School of Criminal Justice at the State University of New York. Professor Hindelang is involved in studies of the utilization of criminal justice statistics, analysis of victimization survey results, and the reliability and validity of self-reported delinquency measures. His publications include *Criminal Victimization in Eight American Cities* (Ballinger, 1975); *Sourcebook of Criminal Justice Statistics* (coauthored with Michael Gottfredson and James Garofalo, Ballinger, in press).

■ EDWARD J. WALSH (Comment on Jenkins and Perrow, ASR April, 1977) is Assistant Professor of Sociology at Pennsylvania State University. His research interests center on the social psychological correlates of occupations at various levels of prestige hierarchy and their implications for stratification theory. His most recent publication is *Dirty Work, Race and Self-Esteem* (Institute of Labor and Industrial Relations, University of Michigan, 1975).

■ JAMES R. LINCOLN (Comment on Grimes, Bonjean, Lyon and Lineberry, ASR August, 1976) is Assistant Professor of Sociology at the University of Southern California. He is involved in the study of strike rates in the U.S. metropolitan areas and organizational structure in Japanese-related firms in southern California. JON OLSON is a Ph.D. candidate in the Department of Sociology at the University of Southern California. Currently he is doing research on the organizational structure of Japanese business organizations.

■ ROBERT M. HAUSER (Comment on Butler, ASR October, 1976) is Professor of Sociology at the University of Wisconsin, Madison. He is continuing his study of the American system of social stratification and the development of socioeconomic inequality over the life cycle. Recently he published (with D. L. Featherman) *The Process of Stratification: Trends and Analyses* (Academic Press, 1977). In 1976 he coauthored (with W. H. Sewell) *Schooling and Achievement in American Society* (Academic Press).

■ JOEL H. LEVINE (Comment on McCann, ASR February, 1977) is Assistant Professor in the Program in Mathematical Social Sciences and the Department of Sociology at Dartmouth College. He is engaged in research on social networks with particular emphasis on interlocking directorates. His other research interests include models of time series, occupational mobility, and multidimensional scaling applied to large data sets with missing data.



NEW PAPERBACKS FROM DELL

WHO SHOULD PLAY GOD? The Artificial Creation of Life and What It Means for the Future of the Human Race

Jeremy Rifkin and Ted Howard. "A startling, readable, up-to-date account of recent biological breakthroughs...the dangers it exposes and the issues it raises are of shattering importance."—ALVIN TOFFLER, author of *Future Shock*. Dell \$1.95.

REDISCOVERY OF THE BODY A Psychosomatic View of Life and Death

Edited, with an Introduction, by Charles A. Garfield, Ph.D. Sixteen authoritative articles on mind-body integration in Eastern and Western disciplines, from Yoga and Aikido to Psychedelics and Rolfing. Laurel \$2.50

NO MORE DYING The Conquest of Aging and the Extension of Human Life

Joel Kurtzman and Phillip Gordon. From bionics and cloning to cold sleep and nuclear-powered hearts, "an important and revolutionary book...hard-edged, well-researched...makes complex scientific discoveries interesting and comprehensible."—*L.A. Free Press*. Laurel \$1.95

BODY, MIND, BEHAVIOR

Maggie Scarf. An explanation of why people behave the way they do. "Lucid, lively...collection of pieces surveying the controversial frontiers of psychology...an intelligent popular introduction to the range of new thinking."—*The Kirkus Reviews*. Laurel \$2.25

Previously published

THE MALE MACHINE

Marc Feigen Fasteau; introduction by Gloria Steinem. "Brilliantly demonstrates the destructiveness of the traditional male value system...informative, perceptive and fascinating."—*Chicago Sun-Times*. Delta \$2.95

TRANSVESTITES & TRANSSEXUALS

Deborah Heller Feinbloom. Who they are, how they feel, what they want and why. "For the nonspecialist, this is the best book there is."—JOHN MONEY, author of *Sexual Signatures*. Modern Psychology Book Club Alternate Selection. Delta \$3.95

MIND AS HEALER, MIND AS SLAYER

A Holistic Approach to Preventing Stress Disorder

Kenneth R. Pelletier, Ph.D. Using self-help techniques and biofeedback, "Pelletier reviews (with an excellent bibliography) the nature of stress, its causes, manifestations, and relation to personality and the many diseases in which it plays a part."—*New England Journal of Medicine. Psychology Today* Book Club Alternate Selection. Delta \$4.95

Dell Publishing Co., Inc.
245 East 47th Street
New York, N.Y. 10017

Dell

How the Kansas tornado of 1966
helps today's students
learn sociology.



GORDON SOCIOLOGY AND AMERICAN SOCIAL ISSUES

Leonard Gordon
Arizona State University
Early 1978 / about 624 pages
Study Guide by Patricia Harvey
Instructor's Manual

Gordon captures student interest by using important social issues — like community organization during disasters, the aged as a minority group, intergroup conflict and accommodation — to explain the principles and methods of sociology. After a presentation of fundamental concepts, institutions, and methods of inquiry, Gordon develops issues dealing with individuals, groups, and the world community, and finally discusses social change and the future.

For each chapter, he provides an introduction, summary, discussion questions, and extensive bibliography, while a glossary concludes the text. A complete instructional package, Gordon offers a Study Guide with detailed learning aids and a substantial Instructor's Manual.

AMERICAN SOCIOLOGICAL REVIEW

EQUALITY, SUCCESS, AND SOCIAL JUSTICE IN ENGLAND AND THE UNITED STATES*

ROBERT V. ROBINSON AND WENDELL BELL

Yale University

American Sociological Review 1978, Vol. 43 (April):125-143

In an exploratory study of matched samples in England and the United States, we construct a path model that explains 26% and 39%, respectively, of the variance in social judgments about the fairness or unfairness of equality. The underdog principle, from which we predict that egalitarians compared to inequality are more likely to be nonwhite, to have low prestige occupations, to have low family incomes, and to identify with the lower and working classes, is accepted. The principle of enlightenment, from which we predict a positive relationship between education and favorable attitudes toward equality, is accepted for England but not for the United States. The principle of an egalitarian Zeitgeist, from which we predict younger people are more egalitarian than older people, is accepted for the United States but not for England. Two additional important causal variables are found. First, a sense of personal equity, that is, a belief that a person has the standard of living that he/she deserves, reduces egalitarian attitudes in England more than in the United States and may reflect a cultural belief that British society is extraordinarily just because social arrangements result from fair rules of the game. While it is of no importance in England, the cultural belief in monetary success reduces egalitarian attitudes in the United States and functions as the belief in the just society does in England.

PURPOSE

Although stratification and inequality have long constituted major foci of sociological inquiry, there has been relatively little sustained empirical study of evaluations of the social justice of equality and inequality in the society as a whole. Yet, as Rainwater (1974:159) notes, "there is no shortage of ready judgment of what is fair, equitable, just and reasonable

or what is unfair, unjust and arbitrary about the rewards and lack of rewards available to different categories of people in society." Rationales for inequality are as ubiquitous as inequality itself, and they are matched in quantity by contrary demands for equality and justifications of them. With a few notable exceptions, such rationales and justifications have largely escaped social scientist's attention.¹

Obviously, conflicts over the justice of equality and inequality have been—and

* This is a revised version of part of a paper read at the annual meeting of the American Sociological Association, New York, 1976. The authors gratefully acknowledge financial assistance from the National Institute of Mental Health through the Yale Comparative Sociology Training Program (Grant No. 5-T01-MH12133) and from the Yale Concilium on International and Area Studies. Also, we thank the Shell Company's Discretionary Grant to the Department of Sociology, Yale University, for supplemental support. Additionally, we thank our colleagues Janet Grigsby, Candace M. Kruttschnitt, H. Wesley Perkins, and Robert H. Ross for their important contributions; Dr. Michael Mann of the London School of Economics and Political Science for his advice in designing the study; and Harry Laracuenta, Bernice A. Pescosolido, and J. L. P. Thompson for their helpful comments on an earlier draft of this paper.

¹ Among the empirical studies that deal with social judgments of equality and inequality in the society at large are Bales and Couch (1969); Bell (1964); Jasso and Rossi (1977); Lane (1959); Mau et al. (1961); Moskos and Bell (1967); Rainwater (1974); Rokeach (1973); Runciman (1966); Seeman (1960); and Turner (1958). Elsewhere, we have evaluated various measurements of attitudes toward equality (Bell and Robinson, 1978). The important work on equity which is based on experimental laboratory studies is relevant to any consideration of evaluations of equality and inequality and we have taken it into account. For example, see Adams (1963); Anderson et al. (1969); Berger et al. (1972); Berkowitz and Walster (1976); Cook (1975); Patchen (1961); and Weick (1966).

still are—at the heart of many struggles over public policies, whether they deal with educational opportunity, job chances, access to housing and public places, or health care. Thus, there is a practical need to know how much equality and inequality for whom, compared to whom, under what conditions, will be justified or tolerated by members of a society. Furthermore, theories of stratification would be augmented importantly by knowing which individuals judge equality or inequality to be just, why they do so, and where they are located in the social structure. That is, knowledge of existing values concerning the fairness of various types and amounts of inequality and how these values are distributed can profitably expand sociological descriptions and explanations of stratification systems.

Thus, our first purpose is to uncover the causes of variations in individual judgments about the fairness of equality or inequality. We propose an explanatory model using path analytic techniques.

Our second purpose is to test the generalizability of the model by its application to data from two societies, England and the United States. These two countries are well suited to our purposes in that much has been written about the justice of equality and inequality in them; and their historical conditions and cultural traditions, although within the same civilizational and language frame, have differed in significant ways regarding inequality. Since Tocqueville's *Democracy in America* was published in the mid-1800s, it has become commonplace to characterize American society as having relatively more equality than European societies and the American people as more highly valuing equality. More recently, Lipset (1963:517) observed that "more than any other modern non-Communist industrial nation, the United States emphasizes achievement, equalitarianism, universalism, and specificity," while Great Britain emphasizes the opposite values of ascription, elitism, particularism, and diffuseness. Similarly, Huber and Form (1973:4) refer to equality, success, and democracy as the three dominant values of the American ideology and Williams (1951:390, 409–11) places achievement

and success and equality among the major value orientations in America.

But complicating the predominant view of American egalitarian values and English inegalitarian values are some other considerations. First, the great stress on achievement and equality of opportunity in America inevitably results in some approval of inequality of condition (Lipset, 1967:369; Kristol, 1968:110). Second, Britain has advanced toward socialism further than the United States and this advancement has been justified in part by egalitarian values (Rose, 1964:39), while, it has been asserted (Foshwald, 1973:38), Americans are "morbidly suspicious of any kind of socialism and social-welfare state." Third, British society, with democratization, industrialization, and urbanization, has been moving away from elitism and toward egalitarianism; and the United States may have become less egalitarian as it adopted a quasi-imperialistic and sometimes reactionary role in international affairs (Lipset, 1968:300). Thus, sorting out the within and between country similarities and differences in social judgments of equality and inequality using data from these two countries should enhance a sociological understanding of such judgments and the range of societal conditions under which they occur.

BASIC DATA

The basic data on which this exploratory study is based are interviews averaging about an hour in length with persons 18 years of age or over in England and the United States. Interviewing in both countries was done during the summer of 1975. In the United States interviews were conducted in the cities of East Haven, New Haven, West Haven, and Woodbridge, Connecticut, which were chosen to provide a sample of individuals representing a wide range of occupational, ethnic and racial backgrounds. Using the *New Haven City Directory, 1974* as our sample frame, we selected every n th name starting with a random number from 1 to n . We fully interviewed 113 persons—or 73.4% of the potential respondents after persons

who had changed address leaving an unoccupied dwelling unit or were too old or ill to be interviewed had been removed.

In London, using statistics for occupation and race provided by the Population Studies Office of the Greater London Council, we chose seven wards which were as comparable demographically to the four New Haven cities as we were able to get. As our Greater London sample frame we used the *Register of Electors* covering all seven wards available in each borough hall. The *Register* included all registered voters age 18 and over. The wards were Kensal Rise, Kenton, Sudbury, and Tokyngton in Brent Borough; East Barnet and Garden Suburb in Barnet Borough; and Ruskin in Southwark Borough. We interviewed a total of 101 respondents in the London area for a 73.2% response rate.

The samples are adequate for the purpose of exploring causal connections within each country. For cross-country differences, of course, gross comparisons cannot be assumed to represent the results of an overall England-United States comparison between representative samples, since neither sample, especially the English one, is representative of the entire country. Nonetheless, country comparisons can be made meaningfully to the extent to which we have matched successfully sampling areas on key demographic variables or have introduced relevant statistical controls.

THE DEPENDENT VARIABLE: EVALUATED EQUALITY

The dependent variable is a measure of the individual's judgment of the fairness of equality or inequality; that is, the individual's value judgment as to the justice of equality or inequality.² Our operational definition is a five-item Index of Evaluated Equality (IEE) which we have previously proposed and tested (Bell and Robinson, 1978).³ The response categories are strongly agree, somewhat

agree, undecided, somewhat disagree, and strongly disagree. The items are:

1. It's fair that rich people who can pay their fines can stay out of jail while poor people may have to go to jail for the same crime.

2. It's unfair that people in some occupations get much more respect from others than people in other occupations.

3. Landlords should be allowed to turn away prospective tenants even if they do so for racial reasons.

4. It would be more fair if people in America (Britain) were paid by how much they need to live decently rather than by the jobs they do.

5. It would be a good thing if Congress and the President (Parliament) decided to distribute all the money in the United States (Britain) equally among all the population.⁴

The respondent must disagree with questions 1 and 3 and agree with questions 2, 4, and 5 to give an egalitarian response, thereby minimizing the effects of acquiescence bias.

The equality or inequality being judged as fair or not by the IEE deals with economic aspects (rich, occupation, pay, and monetary wealth) predominantly, but also social (respect, housing and need to live decently), racial and legal aspects. A factor analysis of the scale resulted in one factor that explains 100% of the common variance of the items in England and a first factor that explains 82% of the variance in the United States, demonstrating the unidimensionality of the scale. Average item-item correlations are about .2 in both countries and the reliabilities are satisfactory, though a bit low. For England, Cronbach's (1951) alpha is .58 and Heise and Bohrnstedt's (1970) omega is .59; for the United States alpha is .56 and omega is .62. Measures of internal or trait validity are also satisfactory; Heise and Bohrnstedt's (1970) validity measure rho is .77 for England and .79 for the United States. Their measure of invalidity, Ψ^2 , is effectively zero for both countries, indicating that there is no variation among the

² In defining justice as fairness, we follow Rawls (1971).

³ For a conceptual clarification of equality and equity, see Bell (1974).

⁴ The alternative wordings used in London are given in parentheses.

index items due to causes or factors other than the one underlying variable.

The response categories for each item are scored zero for "least egalitarian" and four for "most egalitarian," giving a possible range of scores over the five items from zero to twenty. The frequency distributions of total scores on the IEE for England the United States are:

	England (Percent)	United States (Percent)
16 or more (most egalitarian)	11.9%	18.6%
11-15	38.6	40.7
6-10	33.7	31.9
5 or less (least egalitarian)	15.8	8.8
Total	100.0%	100.0%
Number of cases	(101)	(113)

Making a gross comparison, we find Americans only slightly more egalitarian than the English, the mean total score for the IEE in England being 10.2 and in the United States 11.5 (a difference significant at the .025 level). This difference is reflected in a zero-order correlation between country (England = 0, U.S. = 1) and IEE of only .14, with a partial correlation reduced to .08 controlling simultaneously for sex, race and occupation. Although the difference between the countries appears to be genuine, it is small. Thus, as Lipset has cautioned, there may be more similarity than difference between the two countries. Also, responses on the individual items between the two samples are quite similar, except for the one item dealing with race and housing where Americans are clearly more egalitarian than the English, and the five items are identically ranked by the degree of egalitarianism of the responses in the two countries (as listed above from the most to least egalitarian responses).⁵

Furthermore, as we have reported earlier (Bell and Robinson, 1978), more important than the small country differences is the remarkable degree of tolerance of

some inequalities among individuals in both England and the United States. Our results are consistent with the recent finding of Jasso and Rossi (1977:650) that 200 Baltimore city adults would consider an equal distribution of earnings unjust.

AN HYPOTHESIZED CAUSAL MODEL

Our guiding hypotheses for the explanation of variation in evaluations of equality within each country are based on three underlying principles.

1. *The underdog principle.* According to this principle, individuals who objectively benefit from the stratification system in comparison with others are more likely to judge its inequalities to be just. Conversely, people who are objectively less well off are more likely to judge equality to be fair, since greater equality, other things remaining the same, would result in their receiving more societal goods. Thus, our prediction is that underdogs will tend to favor equality. For race, occupation, and family income, our predictions are that persons who are nonwhite,⁶ have relatively low-ranked occupations, and low family incomes are more likely to evaluate equality as fair than are persons who are white, have relatively high-ranked occupations and high family incomes.

We assume that there is also a subjective element involved in the underdog

⁶ Nonwhites in both England and the United States constitute groups subordinated to the white majority, although they differ in composition between the two countries. Eighty-eight percent of the United States nonwhites in the sample were black, 6% were Chinese, and 6% Puerto Rican. The English nonwhites in the sample were 58% black and 42% East Indian or Pakistani. The percentage of nonwhites in each sample slightly underrepresented that in the target communities by about 2% in each country. As regards nativity, 82% of the American nonwhites were born in the United States and 18% born elsewhere, 6% each in Jamaica, Puerto Rico and Taiwan. None of the English nonwhites was born in England, 58% having been born in the West Indies, 25% in Africa, and 17% in India. A comparison of the sample distribution of nativity with that of the target areas shows less than a 2% undersampling of foreign-born in the United States and a 3% oversampling in England. The samples differed somewhat by the length of residence of nonwhites in the communities. In the United States 56%, and in England 67%, of the nonwhites had lived in the area for more than 10 years.

⁵ For a complete evaluation of the Index of Evaluated Equality, including other items not included in the final version of the index, see Bell and Robinson, 1978. Persons not answering all of the items of the IEE were given total scores based on the items that they did answer.

principle, in that individuals may perceive and react to the objective differences in deprivation and privilege. This is to some extent captured by a question on subjective social class, the self-placement of the respondent in one of five social classes. We know from past research that subjective social class is positively related to race, occupation, family income (and education); on the basis of the underdog principle, we further predict that it will be related negatively to favorable attitudes toward equality.

The underdog principle may seem obvious and is, in fact, more or less supported by a number of prior studies (Centers, 1949; McClosky, 1964; Nowak, 1969; Form and Rytina, 1969; Feagin, 1972; and Hamilton, 1972). However, Lane (1959:46-8), in his interviews-in-depth with fifteen American lower-middle and working-class men, found that most of the men had a fear of equality. His respondents thought that the distribution of rewards was just and rational, that under equality "there would no longer be . . . an elite to supervise and take care of people—especially 'me,'" that greater equality "would deprive men of their incentive to work, achieve, and develop their skills," and that it is wrong that "one's own hard earned status [be] given to [the lower orders] as a right and not as a reward for effort." Thus, from Lane's findings, one might decide that the underdog principle is not obvious after all.

2. *The principle of enlightenment.* With enlightenment—by which we mean for England and the United States familiarity with the intellectual achievements of the major values and themes of Western Civilization—comes a greater commitment to the idea of equality as a positive value. The history of democratic revolutions—and evolutions—implies equality in a variety of ways, remains a charter myth of these societies, and is taught the young as sacred past. We measure enlightenment by amount of education.

Of course, there is at least one aspect other than enlightenment to education and that is socioeconomic status. But with occupation and family income separately measured and controlled, the enlighten-

ment aspect of education, if it exists, should be revealed in our analysis, with well-educated people being more likely to favor equality as fair than less educated people.

An alternative to our principle of enlightenment may be found in Turner's (1960) observation that the English and American educational systems differ both in their structure and in what is taught. Under the English system, characterized by sponsored mobility, recruits to higher education are chosen by the existing elite. The ideals inculcated by the system support the view that persons of inferior status should resign themselves to their fate, while elite recruits should accept their roles as tomorrow's leaders. The American system, by contrast, is characterized by contest mobility, whereby higher education may be won by the student as a reward for effort and intelligence. The values taught are almost exactly opposite those taught in England—that anyone, no matter what his/her origins, can advance if he/she is willing to try hard enough. If one assumes that socialization into these values occurs at a very early age, then there is no reason to draw any conclusion from Turner's analysis about the effect of increasing education beyond a certain point on egalitarian attitudes. If, however, one assumes that such socialization occurs throughout the educational ladder, then one might well predict from Turner's analysis that persons with more education in America should be more egalitarian while in England they should be more inegalitarian, the opposite of our prediction from the enlightenment principle.

3. *The principle of an historical shift toward an egalitarian Zeitgeist.* It is elementary to point out that in most societies throughout human development inequalities have existed and have been accepted as normal and legitimate. Yet in spite of the

great diversity of patterns, certain trends seem to emerge from a consideration of the last hundred and fifty years. Everywhere there seems to have come about a steady erosion in the legitimacy accorded to social inequality. If social inequality continues to exist as a fact, it is no longer accepted by all

as a part of the natural order but is challenged, or at least questioned, at every point. (Beteille, 1969:366)

In England, for example, Rose (1974: 51-2) observed that since the enacting of the three "major landmarks of the mixed-economy welfare state"—the Beveridge Report on Social Welfare, the Keynesian Full Employment White Paper of 1944, and the Butler Education Act of 1944—"Conservative and Labour governments have sought to combine economic prosperity and a generous provision of public services for the electorate." The last twenty years of social welfare legislation was merely part of a trend extending back at least a century, with the provision of public services having expanded under governments of each political shade (Rose, 1974:129-30).⁷

In the United States the early 1960s saw the "rediscovery of poverty" in academic and official circles, partly due to the influential publications of Harrington (1962) and MacDonald (1963). Perhaps as a direct result, "the first major social legislation (Civil Rights Act, the Economic Opportunity Act, the Manpower Development and Training Act) in over a quarter of a century was enacted" (Miller and Rein, 1966:429). The War on Poverty and the Great Society it was to establish became familiar catch phrases. The decade and a half since the rediscovery of poverty has seen an increase in demands for equality by blacks, poor whites, college students, women, Chicanos, and American Indians (Huber and Form, 1973:7), a trend which has led Gans (1973) to claim that events of the 1960s and 1970s constitute an egalitarian revolution.

Thus, because in both England and the United States there has been a secular trend toward an increase in favorable attitudes toward equality, we predict that younger persons will be more likely than older persons to judge equality to be fair, each group reflecting a generational difference based on the dominant values of its most significant period of political socialization.

⁷ This is not to say, of course, that rising levels of living have resulted in economic equality, as Westergaard and Resler (1975:118) point out.

This hypothesis runs directly counter to what we would predict using the underdog principle: assuming that older adults are more disadvantaged than younger adults, we would hypothesize that age would be positively correlated with the IEE.

Since we do not have cohort data, we cannot distinguish, however, between our hypothesis, as derived from the Zeitgeist principle, and still another alternative hypothesis, that older age produces more conservative responses, that is, life-cycle effects, because from each we would expect the same finding, a negative correlation between age and egalitarian attitudes. Our data do permit us to see whether the underdog principle on the one hand or either the Zeitgeist or life-cycle hypotheses on the other hand is a better fit to the age data.

In Figure 1 we summarize the above hypotheses as formulated from our three principles.⁸ We make little of the predicted relationships between race, education, occupation, and family income with subjective social class, since these relationships are fairly well established from past research. Thus, the key predicted relationships are the six for each country—or twelve in all—involving evaluated equality.

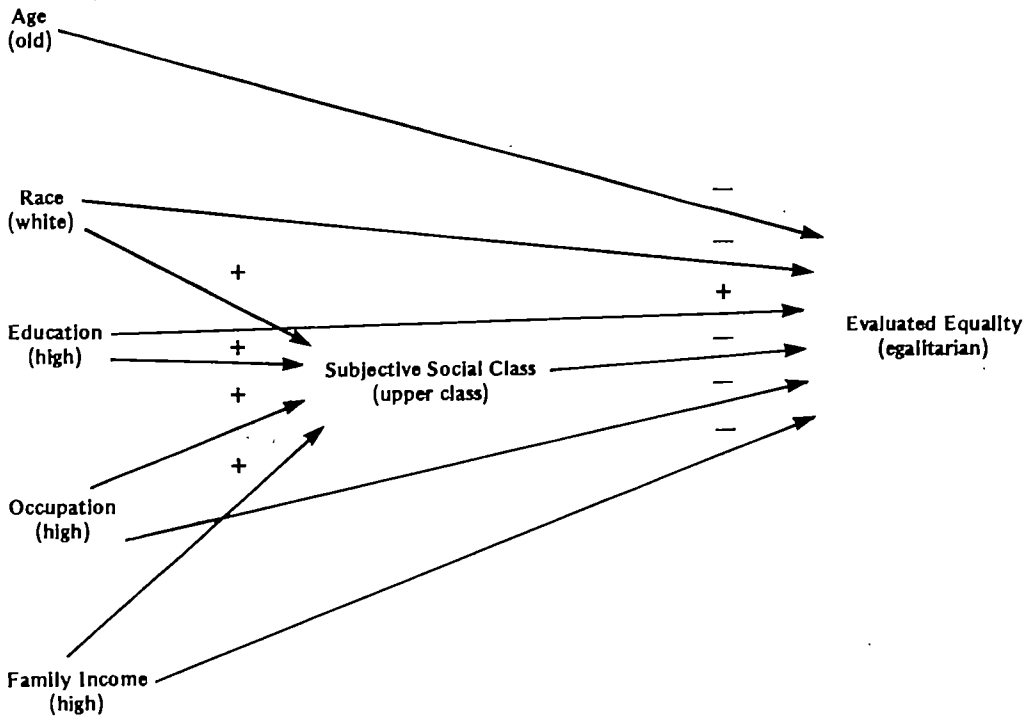
ZERO-ORDER RELATIONSHIPS

From the correlation matrix given in Table 1,⁹ note the correlations of each of the six independent variables with evaluations of equality.¹⁰ For England, the signs

⁸ We make no assumptions of causality among age, race, education, occupation, and family income, since such a model would be more appropriate to the study of the process of status attainment than to our purpose here. The correlations among them are shown in Table 1. We do assume that subjective social class occurs in time after the other five independent variables.

⁹ Missing data on all variables are recoded to the mean of the variable in question for the correlation and regression analyses.

¹⁰ Measurement of the independent variables is as follows: age is straightforwardly chronological age as reported by the respondent. Race is dichotomized into white and nonwhite; education is coded in comparable categories in England and the United States with advice of Dr. Michael Mann of L. S. E. Categories are (with the United States in parentheses): graduate degrees (graduate or pro-



* The high pole of each variable is given in parentheses.

Figure 1. Hypothesized Paths in a Causal Model Explaining Evaluated Equality *

of all six of those correlations are in the predicted direction, and the correlations for race, occupation, and subjective social class are statistically significant ($p <$

.05).¹¹ For the United States, five out of six of the signs are in the predicted direction, the exception being education, and the correlations for age, race, and subjective social class are statistically significant

fessional training); college degree (college degree); G.C.E. Advanced Level, Higher Level Diplomas from further educational institutions, C.S.E. H Level, accounting qualifications (Partial college—at least one year—or specialized training); G.C.E. Ordinary Level, Ordinary Diplomas, City and Guilds qualifications, C.S.E. O Level (high school graduate); no qualifications, left school age 15 or over (10–11 years of school); no qualifications, left school age 12–14 (7–9 years of school); and no qualifications, left school below age 12 (less than 7 years of school). Occupation is coded in Treiman's (1977) prestige units which are specifically designed for cross-national purposes and have been previously used in a comparison of England and the United States (Treiman and Terrell, 1975). Housewives who were not employed outside the home are assigned their husband's occupation and young people listing their parents as chief breadwinners in the household are given their father's occupation. Family income is the respondent's best estimate of all the income earned or received by family members before taxes in the respondent's household.

Subjective social class is measured by both an open- and closed-ended question.

We obtained the following distribution of responses:

	England (Percent)	United States (Percent)
Upper class	0%	1%
Upper-middle class	7	9
Middle class	48	42
Working class	43	40
Lower class	2	8
Total	100%	100%
Number of cases	(90)	(107)

¹¹ The correlations reported in this paper, uncorrected for attenuation due to measurement error (e.g., Bohrnstedt and Carter, 1971), are conservative estimates of the true correlations. This is especially true of the correlations with the IEE since its reliability is somewhat low (about .60). If corrected for attenuation, these would be about a third again as large as those reported here. This does not affect the country comparison since the reliability of the IEE is virtually the same in England and the United States. Since our exploratory samples are small, we have adopted the conservative procedure of reporting only the smaller, uncorrected correlations, so that only relationships which are strongly supported by the data will be statistically significant.

Table 1. Correlation Matrix, Means, and Standard Deviations for Variables in the Causal Model, England and the United States ^a

Variables									
Variables	Age	Race	Education	Occupation	Family Income	Sense of Personal Equity	Subjective Social Class	Perceived Monetary Success	Evaluated Equality
Age (old) ^b	—	.11	-.48°	-.16	-.33°	.08	-.09	-.30°	-.13
Race (white)	.26°	—	-.09	.14	.11	.22°	.22°	-.01	-.32°
Education (high)	-.27°	.18°	—	.55°	.33°	.10	.36°	.30°	.05
Occupation (high)	-.01	.23°	.58°	—	.27°	.27°	.52°	.35°	-.29°
Family Income (high)	-.19°	.11	.20°	.33°	—	.09	.34°	.30°	-.02
Sense of Personal Equity (high)	.21°	.50°	.29°	.24°	.18°	—	.22°	.13	-.31°
Subjective Social Class (upper class)	.09	.11	.30°	.40°	.24°	.27°	—	.31°	-.24°
Perceived Monetary Success (high)	-.15	-.03	.24°	.10	.16°	.03	.20°	—	.03
Evaluated Equality (egalitarian)	-.39°	-.50°	-.07	-.09	-.08	-.38°	-.20°	-.15	—
(U.S. \$00)									
Means									
England	45.17	1.88	3.35	46.48	100	2.46	2.60	2.70	10.25
U.S.	45.20	1.84	4.15	44.44	139	2.33	2.56	3.06	11.49
Standard Deviations									
England	17.69	.33	1.34	15.32	60	.91	.64	1.59	4.39
U.S.	17.29	.37	1.57	14.95	73	.93	.78	1.44	4.34

^a Correlations for England are above the diagonal and for the U.S. are below.^b The high pole of each variable is given in parentheses.* Significant at $p < .05$.

($p < .05$). None of the correlations whose sign is in the direction opposite to our prediction is significant in either country.

Also from Table 1, the correlations of race, education, occupation, and family income with subjective social class are all positive in each country, as expected, and seven out of the eight are statistically significant ($p < .05$), the only failure being the correlation with race in the United States, which is .11. In each country, the highest of these correlations is that between subjective social class and occupation, .52 in England and .40 in the United States.

Of greater interest than the zero-order correlations, of course, are the net effects of the variables. Before presenting these findings, however, we wish to introduce two additional variables.

SENSE OF PERSONAL EQUITY AND PERCEIVED ACHIEVEMENT OF MONETARY SUCCESS

We explored the effects of a number of different additional variables in the model. Two were kept in the final analysis, because they made the most sense theoretically, related to an important aspect of British or American culture, and explained some additional part of the variance in evaluated equality.

The first is a measure of the individual's sense of personal equity. We asked our respondents, "In general, do you personally have the standard of living that you feel you deserve?" We were interested in their evaluations of the fairness or justice of their personal levels of living. Their responses were:

	England (Percent)	United States (Percent)
No, I deserve less	3%	2%
Yes, I have what I deserve	66	60
I'm not sure	4	8
No, I deserve more	27	30
Total	100%	100%
Number of cases	(98)	(112)

Although the English respondents were slightly more likely than the Americans to feel that their standard of living is equitable or fair, the two samples are remarkably similar on this variable. In both sam-

ples, most respondents think that they deserve what they have. Yet a sizable minority (27% in England and 30% in the United States) feel that they deserve more, indicating that considerable dissatisfaction exists in both countries.

The second additional variable that we introduce into the model is the individual's perceived achievement of monetary success. Each respondent was asked the following:

I'm going to read off a few things and, as I do, please tell me how close you have come—or think you *will* come—to getting each in your own lifetime.

Respondents were then asked to select a response from a card that was handed to them and twelve items, or life goals, were read.¹² The item we used to measure perceived achievement of monetary success, the response categories, and responses from England the United States are given below:

"Really getting ahead by earning a lot of money."

	England (Percent)	United States (Percent)
I will definitely get this; or		
I already have this	18%	20%
I may get this	26	31
Undecided; or don't know	1	1
I may never get this	20	30
I won't get this	35	18
Total	100%	100%
Number of cases	(101)	(109)

Americans were only about half as likely as the English to say that they won't achieve monetary success and were slightly more likely to say that they might achieve such success. The difference in perceived monetary success between the two countries is significant ($p < .05$). Although the American sample is somewhat better off monetarily than the English (the mean family income in the American sample was US\$13,900 but only US\$10,030 in the English sample), we suggest later that the higher level of living in America may

¹² We adapted this item from the Index of Perceived Achievement of Cultural Life Goals constructed by Meier (1962).



not influence perceived achievement of monetary success as much as a general cultural belief does.

Although we had no hypotheses regarding the relationship of either of these variables to evaluated equality before we began the data analysis, we thought the underdog principle should apply. That is, individuals who think they have the standards of living they deserve, compared with people who don't, should be less likely to be favorable to equality. Similarly, people who perceive themselves as monetarily successful—or as likely to be so in the future—should be less likely to favor egalitarianism than people who perceive themselves as not being so successful.

The zero-order correlations of personal equity with evaluated equality shown in Table 1, bear this out in both England ($-.31$) and the United States ($-.38$). But perceived monetary success is negatively correlated with the IEE only in the United States ($-.15$). In England, the correlation is practically zero ($.03$). But, as we shall see, each may tell an additional story as well.

A PATH MODEL

Following Alwin and Hauser (1975), we present the decomposition of effects in the path model for England and the United States in Table 2.¹³ Since our chief concern is with the explanation of judgments about the justice of equality, we point out only a few highlights of the findings dealing with the other dependent variables.

First, the multiple correlation of social background variables with the sense of personal equity is higher in the United States ($.58$) than in England ($.34$). The individual's sense of personal equity (except insofar as it results from occupational prestige) is less grounded in objective position in England than it is in the United States. This is so, we suggest, because of a societal belief, perhaps now on the wane, that England is a just society.

The belief of the English that their society is just by virtue of a general agreement

on and conformity with norms of social conduct has been mentioned in many accounts of English national character. Anthony Glyn (1970:205), for example, says that the "idea of fairness stems from the fundamental British assumption that there are rules for everything in life. Should someone somewhere transgress or ignore these rules, a cry of 'unfair' will go up."

Maillaud (1946:38) says, "Everywhere in England there is a strong sense of what is 'done' and 'not done,' 'proper' and 'improper,' and, above all, 'right' and 'not right,' both as regards trifling matters and issues of great moment." And Commager (1974:752) in his analysis of English traits agrees, saying that the "English have a highly developed sense of justice and right. . . ."

Survey data support these impressions. In their study of images of the United States and Britain among Indians who had been trained in the West, for example, Useem and Useem (1954:74) found that phrases used to portray the personal integrity of the British prominently included "fair" while the top four qualities used to portray Americans did not. Rose and Kavanagh (1976:560) report that comments about the British system of government among their Scottish respondents clearly and prominently included "fairness" and "justness." And in our English sample, 20% of the respondents volunteered "justice" or "fairness" when asked, "What are the things about this country of which you are most proud?" Only 5% of the American respondents gave such answers.

Such a general belief in a just society, we assume, leads to satisfaction with what one has; that is, to a sense of personal equity, and, as we report below, to the judgment that, under the prevailing conditions of British society, inequalities are fair. This is not a tautology, since, under different conditions, the belief in the just society could conceivably support egalitarian social arrangements.¹⁴

¹³ For an evaluation of the use of ordinal measures in path analysis, as we have done here, see Labovitz (1967; 1970; 1971) and Boyle (1971).

¹⁴ Our variable, the sense of *personal equity*, measures only indirectly the belief of the just society. It may have both an individual component and a societal one. That is, an individual may feel he/she has what he/she deserves quite independently of judgments about the justness of society as a whole,

Table 2. Decomposition of Effects in a Model Explaining Evaluated Equality, England and the United States

Dependent Variable	Predetermined Variables	England				United States			
		Total Effect	Indirect Effect	Direct Effect	Multiple R	Total Effect	Indirect Effect	Direct Effect	Multiple R
Sense of Personal Equity (high)	Age (old) *	.13	—	.13		.21	—	.21*	
	Race (white)	.17	—	.17*		.39	—	.39*	
	Education (high)	.04	—	.04		.28	—	.28*	
	Occupation (high)	.23	—	.23*		-.05	—	-.05	
Subjective Social Class (upper class)	Family Income (high)	.04	—	.04	.34	.14	—	.14*	.58
	Age	.08	—	.08		.17	—	.17*	
	Race	.15	—	.15*		-.04	—	-.04	
	Education	.14	—	.14		.17	—	.17	
Perceived Monetary Success (high)	Occupation	.38	—	.38*		.25	—	.25*	
	Family Income	.21	—	.21*	.59	.16	—	.16*	.45
	Age	-.21	—	-.21*		-.04	—	-.04	
	Race	-.05	—	-.05		-.07	—	-.07	
Evaluated Equality (egalitarian)	Education	-.02	—	-.02		.26	—	.26*	
	Occupation	.29	—	.29*		-.08	—	-.08	
	Family Income	.16	—	.16	.46	.14	—	.14	.29
	Age	-.06	-.06	.00		-.34	-.03	-.31*	
Subjective Social Class (lower class)	Race	-.25	-.06	-.19*		-.41	-.03	-.38*	
	Education	.19	-.03	.22*		-.14	-.09	-.05	
	Occupation	-.38	-.05	-.33*		.12	.00	.12	
	Family Income	.02	-.01	.03		-.12	-.06	-.06	
	Personal Equity	-.20	—	-.20*		-.10	—	-.10	
	Subjective Social Class	-.11	—	-.11		-.09	—	-.09	
	Perceived Monetary Success	.13	—	.13	.51	-.18	—	-.18*	.63

* The high pole of each variable is given in parentheses.

* Direct effect is significant at $p < .05$, one-tailed test.

Second, with respect to perceived achievement of monetary success, the social background variables explain considerably more of the variance in England (21%) than in the United States (8%), just the opposite to the case of a sense of personal equity. We confront here a cultural belief of American society, the American Dream of monetary success—"rags to riches"—about which so much has been written that no special case need be made here.¹⁵ Our data show that Americans, more than the English, believe in it to some extent regardless of whether or not they are disadvantaged. Kerckhoff (1977:570-1) reported results that support this. He found that English boys' expectations were far more realistic than those of American boys in that they were more closely tied to social origins.

Third, subjective social class is more highly related to objective measures of socioeconomic position in England than in the United States. This is true, for example, of occupation and income, though not education. This finding is consistent with Lipset's (1966:163) claim that the Ameri-

can lower class exhibits relatively little class consciousness, with Alford's (1963:109) observation that the working class in Great Britain is class-conscious while the American working class is not, and with a recent study of class consciousness in Britain and the United States (Robinson and Kelley, 1977).

Fourth, there are some interesting country differences in individual variables not adequately accounted for by the above general statements. In the case of personal equity, subjective class status, and perceived monetary success, for example, occupation is clearly the key explanatory variable in England. In the United States, however, this is so only in the case of subjective social class. In explaining personal equity, we find in the United States that race is the key variable, followed by education. In fact, occupation has the smallest correlation with personal equity and its sign is minus. In explaining perceived monetary success among Americans, we find education, not occupation, is the key variable.

Fifth, and finally, we turn to the relationships with evaluated equality. As we hypothesized, the total effect of age on evaluated equality is negative and significant in the United States (-.34). But, though negative, it is tiny (-.06) in England. The principle of an egalitarian *Zeitgeist* receives no support for England, but it is consistent with the facts in the United States, where age is one of the most important determinants of egalitarian attitudes.

Race is a determinant of evaluated equality in England and is even more important in the United States.¹⁶ Nonwhites, with other variables controlled, are more likely than whites to judge equality to be fair, as we predicted from the underdog principle. In England the total effect of race is -.25, with about a fourth of this

or the two beliefs may be highly linked together, perhaps causally. As a test of the latter assumption, we considered the responses to another question, "In general, do you think that the political system of this country works for the interest of most of the people most of the time?" On this question, which we use as a measure of the sense of *societal equity*, the English were, in fact, more likely than Americans to respond that the system generally worked for the interest of most of the people: 32% of the English respondents thought the system definitely worked on everyone's behalf and another 35% thought it did so to some extent, contrasted with only 18% and 37% respectively, of the Americans ($p < .01$ level). More striking than these differences is the remarkable lack of variation among English respondents across categories of major social stratification, only 1% of the variance in sense of societal equity being explained by age, race, education, occupation and family income in England, compared with 12% in the United States. Although the zero-order correlation between personal and societal equity is higher in the United States (.25) than in England (.19), when we control for age, sex, race, education, and occupation, the direct effect in the United States is reduced to .08 while in England it holds to .18.

¹⁵ For example, see Chinoy (1955). Our findings are not necessarily incompatible with the relative deprivation perspective in that belief in the rags-to-riches myth may involve considerable variation in individual definitions both of rags and riches. See Hyman (1953) and Runciman (1966).

¹⁶ Since here for evaluated equality, and earlier for personal equity, race is found to be a more important determinant in the United States than in England, we examined effects of the differential racial composition of our samples. Within the nonwhite group in each country, there is considerable homogeneity of responses on both of these items even when country of origin, length of residence in the community, and generation are taken into account.

being indirect, while in the United States the total effect of race is $-.41$ with almost all of the effect being direct.

Education, as we predicted from the enlightenment principle, is positively correlated with evaluated equality in England. There, well-educated persons are more likely to be egalitarian than are the less well-educated. In the United States, the status rather than the enlightenment aspect of education appears to be dominant; the total effect of education is negative and about two-thirds of it is indirect. Well-educated people in the United States tend to be slightly inegalitarian in their own right, but are even more inegalitarian because they are satisfied with their standards of living, have higher social class identifications, and believe they had or would achieve monetary success. Our finding, on the surface, contradicts inferences one might draw from Turner's emphasis on the inegalitarian nature—both in structure and content—of the English educational system relative to the American system.¹⁷

Occupation, with a total effect of $-.38$, is the most important determinant of egalitarian attitudes in England, lower-status persons, as predicted, being more likely than higher-status persons to judge equality to be fair. In the United States, however, the total effect is much lower and in the opposite direction ($.12$); with other variables controlled, persons of lower occupational status in the United States are somewhat more likely to be inegalitarian than are higher-status persons. In England there is a small indirect effect through the three subjective measures;

that is, persons with high-status jobs are more inegalitarian in their own right but are even more so because they tend to identify with the middle or upper classes, to feel that they deserve their privileged position, and to perceive themselves as successful. In the United States, however, all of occupation's total effect is direct.

The results in England support the underdog principle and cast doubt on the embourgeoisement theory that the increasing affluence of the English working class is causing them to become more conservative, a theory that Goldthorpe and Lockwood (1963; Goldthorpe et al., 1969) have also rejected. Taken together, our results require qualification of Alford's (1963:109) observation that "British workers . . . are class conscious but accept the leadership of the aristocracy and the well-educated. . . . American workers are not particularly class conscious but are equalitarian." Although our data suggest that the English are more class-conscious in the sense of having an awareness of belonging to a particular class than Americans and that the English (both working and middle classes) are less egalitarian than the Americans, they also show that the English working class is more egalitarian than the English middle class, while the American working class is actually less egalitarian than the American middle class. If in the United States people do, in fact, place a greater emphasis on consumption and on the availability of opportunity than people in England—and many writers including Lipset (1963:517) concur—then the hypotheses, proposed by Lane (1959:49–50), that lower-status groups would tend to denigrate those below them and that there would be more intraclass than interclass rivalry may explain our findings.

The individual's sense of personal equity is negatively related to evaluated equality in both countries; that is, controlling for other variables in the model, people who feel that they deserve a better standard of living are more likely to judge equality to be fair than people who are satisfied. This is stronger in England where the total (and direct) effect of personal equity is $-.20$ than in the United States where it is $-.10$. Given the

¹⁷ In their social survey carried out in Glasgow, Rose and Kavanagh (1976:554) find that education is negatively related to traditional values in Scotland. For example, less educated people were more likely to think that a monarchy is needed and to feel pride in the Queen than more educated respondents. In our English sample, education is also negatively correlated with pride in the British tradition and heritage and with pride in the Queen. These variables, in turn, correlated $-.24$ and $-.18$ with the IEE. Decomposing the effects, we find that less educated persons in England were more inegalitarian than more educated persons mainly because they were more committed to traditional values: three-fourths of the total effect of education on egalitarian attitudes is indirect through pride in the British tradition.

additional fact that the sense of personal equity is somewhat freer from social structural variables in England than in the United States, a belief in the justice of one's own position in England is a more important source of free-floating cultural support for inequality than in the United States. This finding seems consistent with the tendency of the English to be somewhat more deferential to their leaders and more accepting of their social status than others as noted by several observers (Alford, 1963:109; McKenzie and Silver, 1968: 163ff; Nordlinger, 1967:14ff; Turner, 1960).¹⁸

In England, persons who perceive themselves as having achieved or being likely to achieve monetary success are somewhat more likely to be egalitarian than those who see themselves achieving less; although the relationship is not significant, it may be a shadowy remnant of past noblesse oblige. In the United States, perceived monetary success is significantly and negatively related to favorable evaluations of equality (-.18). That is, in the United States with other variables—including family income—controlled, people who perceived themselves as achieving monetary success are less likely to judge equality to be fair than those who perceive themselves as failing in such achievement. Thus the belief that the American Dream will work for oneself, *largely apart from one's current position*, is a factor working against the ideal of equality of condition. In America, some persons who might have been egalitarian by virtue of their disadvantaged position in society, believe the American Dream that they might one day be successful, and as a result, judge inequality to be fair since they can imagine themselves on the privileged side of an unequal distribution. By the same token, some Americans who

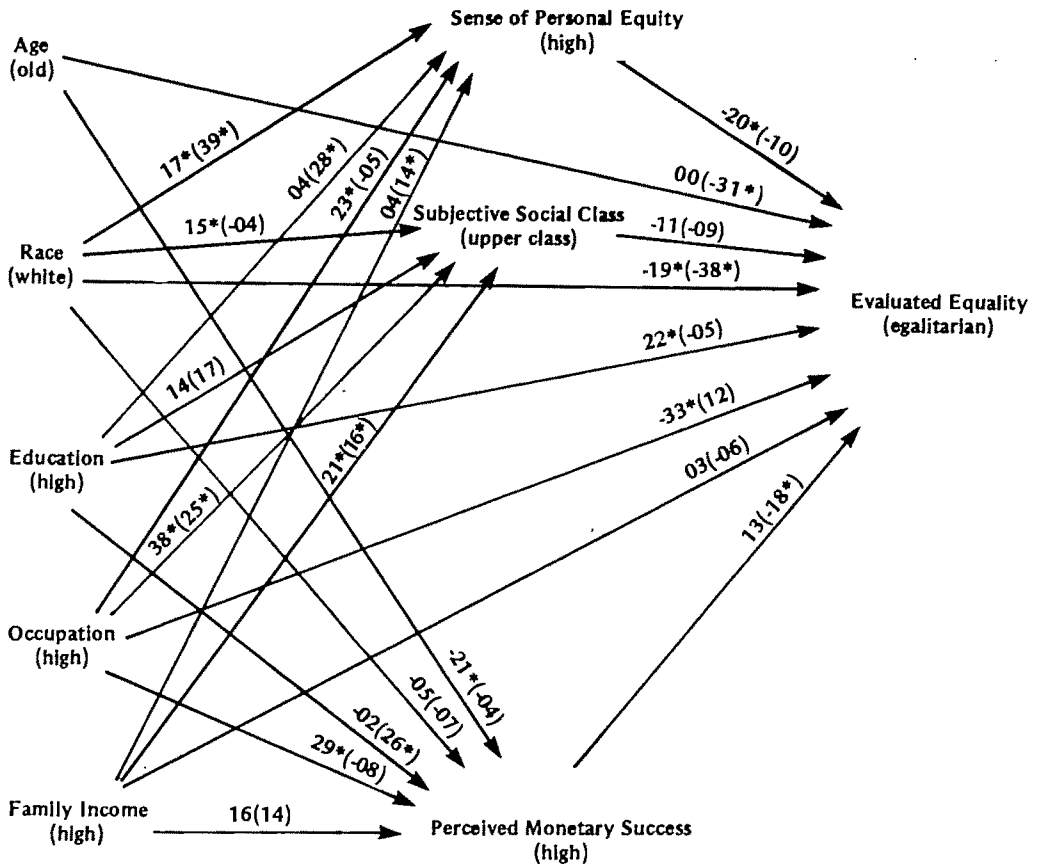
are objectively well off, but who have not bought the American Dream, are egalitarian, despite the fact that their objective interest may lie in continued inequality.¹⁹

Overall, the model explains enough of the variance in each country to be taken seriously. The social background variables (i.e., age, race, education, occupational status, and family income) alone explain 21% and 35% of the variance in evaluated equality, respectively, in England and the United States. This finding contrasts with that of Jasso and Rossi (1977:649) who found that less than 1% of the variance in judgments about the fairness of income inequality was explained by background variables. When we add subjective social class, sense of personal equity, and perceived monetary success we explain 26% of the variance in evaluated equality in England and 39% in the United States. The path model is given in Figure 2.

With other variables controlled, eight of the twelve hypotheses derived from the three principles are supported by the signs of the correlations, six out of eight for the underdog principle and one out of two in each case for the enlightenment and Zeitgeist principles. No correlation whose sign is opposite to that predicted is significantly larger than zero, while five correlations with signs as predicted are significant. Thus, there is more support in favor of than opposed to the underdog principle, but the enlightenment principle does not seem to apply to the United States where the socioeconomic status aspects of education seem dominant and the Zeitgeist principle may not apply to England where young persons, other things being equal,

¹⁸ Our data provide further corroboration that the English tend to be more deferential than Americans. When we asked our respondents, "Do you feel that on most important issues the national leaders of this country know better than the average voter what should be done?", 20% of the Americans answered "definitely yes" compared to 33% of the English. The difference between countries is significant ($p < .01$).

¹⁹ We tested whether the differences between countries in the effects of each of the causal variables on evaluated equality are statistically significant. Using a combined sample of all the cases from both England and the United States and adding a dummy variable for country and interaction terms for each variable with country, we regressed evaluated equality on all the variables reported in Table 2. The interactions of country with four variables—age, education, occupation and perceived monetary success—are significant ($p < .05$), indicating that there are significant differences between England and the United States in the effect these variables have on evaluated equality.



* Decimals omitted; selected paths only.

* Path is significant at $p < .05$, one-tailed test.

Figure 2. Path Model Explaining Evaluated Equality in England and the United States (in Parentheses) *

are no more egalitarian than are older persons.

Note from Table 2 that, for England, occupation remains the most important determinant of evaluated equality, and that education, personal equity, and race also have significant effects, with persons who have low prestige occupations, who have higher educations, who see themselves as deserving a higher standard of living, and who are nonwhite being most likely to judge equality to be fair. For the United States, race is the most important determinant closely followed by age, while the only other significant effect is produced by perceived monetary success, with persons who are nonwhite, are younger, and believe that they have not or will not achieve monetary success being most likely to judge equality to be fair.

CONCLUSION

From a 1975 exploratory study, we proposed a causal model to explain individual variations in judgments concerning the fairness of equality or inequality based upon a new Index of Evaluated Equality. We evaluated the generalizability of the model in a preliminary way by its comparative application to two countries, England and the United States.

Our guiding hypotheses were derived from three principles: (1) the underdog principle, from which we predicted that egalitarians will be found more frequently among nonwhites than whites, more among persons of low prestige occupations than high prestige occupations, more among persons with low family incomes than high family incomes, and more

among people with lower and working class identifications than middle and upper class identifications. The signs of six out of eight of these predictions are confirmed with other variables controlled, but the predictions involving family income in England and occupation in the United States are not. Overall, we conclude that there is an underdog effect contributing to favorable attitudes toward the justice of equality.

(2) The principle of enlightenment, from which we predicted positive relationships between education and evaluated equality, is accepted for England but not for the United States. In fact, the relationship in the United States, though small, is in the other direction, suggesting that the socioeconomic status aspect of education, rather than the enlightenment aspect, is at work along with the underdog principle. These findings are of added significance since they seem to run counter to the widely held idea that British education inculcates elitist values compared with the more egalitarian values of American education. Of course, the American devotion to equality of opportunity and contest mobility might produce people who believe that they fairly earned what they've got and that the resulting inequality was, therefore, fair.

(3) The principle of an historical shift toward an egalitarian *Zeitgeist* is accepted for the United States, where younger people are considerably more likely to be egalitarian than older people, but not for England, where there is no difference with other variables controlled. It is accepted even for the United States, however, only with caution, because the findings are consistent with a competing hypothesis, that conservatism increases with age, and the lack of cohort data make any conclusion shaky. It seems sound, however, to rule out in the United States any gross effects from the underdog principle from which older people would be viewed as disadvantaged and therefore more favorable to equality than younger people.

In the path analysis, occupation is consistently a key explanatory variable in England, and has the highest correlation with sense of personal equity, subjective social class, perceived monetary success,

and evaluated equality. In the United States, to the contrary, occupation is the most important independent variable only in the case of subjective social class; education is most important in explaining perceived monetary success, while race is most important in explaining a sense of personal equity and evaluated equality.

In the course of the analysis, we made a number of discoveries having to do with differences between England and the United States. From an earlier report (Bell and Robinson, 1978), we concluded that, although Americans have slightly more favorable attitudes towards equality than the English, many people in both countries—about 50% in England and 41% in the U.S.—judge inequality to be just. For England, we now see that part of the explanation may be found in a sense of personal equity, a belief by a person that he/she has the standard of living that he/she deserves. This may reflect a cultural belief that England is a just society and may function to reduce egalitarian attitudes. The fact that a sense of personal equity is somewhat less correlated with social background factors and is twice as highly correlated (negatively) with egalitarian attitudes in England than the United States lends some additional support to this interpretation.

In the United States, the cultural belief in monetary success equally open to all appears to serve the same function that the belief in the just society does in England. Americans are somewhat more likely than the English to believe that they have or will "really get ahead by earning a lot of money" and in the United States such a belief is negatively correlated with favorable attitudes toward the justice of equality. Not so in England.²⁰

²⁰ For the United States, we suggest that the belief in the American Dream of monetary success to some extent serves real people in real situations just as Rawls's (1971) hypothetical rational choosers are served by the original position and the veil of ignorance as they decide upon a social contract. Thus, if people withhold judgment about where they are going to end up in the stratification system, they are similar to being in the original position since they have covered their social position with a veil of ignorance. They may favor inequality, perhaps throughout their lifetimes, even if they continue to be disadvantaged in the present, in the hope that sometime in the future they will become privileged by the system.

Thus, our findings may reflect the important cultural difference that belief in the just society in England and belief in monetary success in the United States function to produce acceptance of inequality, with each belief less closely linked to social background factors in the country in which it operates than in the other country.

These beliefs are consistent with Lipset's and Turner's earlier observations on the English and American value systems. Lipset (1963:517) claimed that the English value system emphasizes ascription, while in America achievement is more highly valued. Turner (1960) observed that the English are taught at an early age to reconcile themselves to their lot in life, while Americans are taught that anyone, no matter how humble his/her origins, can achieve success through hard work.

From our data, however, what is crucial is not the modest difference between the two countries in how widely held these beliefs are, but the difference in how they differentially function to support inegalitarian attitudes. Our findings support the generalization that every society, in the face of its particular historical contingencies, provides a rationale, myth or belief, that enables its members to cope with their position in the stratification system. Such a rationale invites people to accept and condone existing inequality as generally just and reasonable. The invitation, however, is not always accepted, as our data show, by young, enlightened or, especially, underprivileged members of society.

REFERENCES

- Adams, J. Stacey
1963 "Toward an understanding of inequity." *Journal of Abnormal Psychology* 67:422-36.
- Alford, Robert R.
1963 *Party and Society: The Anglo-American Democracies*. Chicago: Rand McNally.
- Alwin, Duane F., and Robert M. Hauser
1975 "The decomposition of effects in path analysis." *American Sociological Review* 40:37-47.
- Anderson, Bo, Joseph Berger, Morris Zelditch, Jr. and Bernard P. Cohen
1969 "Reactions to inequity." *Acta Sociologica* 12:1-12.
- Bales, Robert F., and Arthur S. Couch
1969 "The value profile: a factor analytic study of value statements." *Sociological Inquiry* 39:3-17.
- Bell, Wendell
1964 *Jamaican Leaders: Political Attitudes in a New Nation*. Berkeley: University of California Press.
1974 "A conceptual analysis of equality and equity in evolutionary perspective." *American Behavioral Scientist* 18:8-35.
- Bell, Wendell, and Robert V. Robinson
1978 "An index of evaluated equality: measuring conceptions of social justice in England and the United States." In Richard F. Tomasson (ed.), *Comparative Studies in Sociology: An Annual Compilation of Research*, Vol. 1. Greenwich, Ct.: JAI Press. In press.
- Berger, Joseph, Morris Zelditch, Jr. and Bo Anderson (eds.)
1972 *Sociological Theories in Progress*, Vol. 2. Boston: Houghton Mifflin.
- Berkowitz, Leonard, and Elaine Walster (eds.)
1976 *Equity Theory: Toward a General Theory of Social Interaction*. *Advances in Experimental Social Psychology*, Vol. 9. New York: Academic Press.
- Béteille, André
1969 "The decline of social inequality?" Pp. 362-80 in André Béteille (ed.), *Social Inequality: Selected Readings*. Baltimore: Penguin.
- Bohrnstedt, George W., and T. Michael Carter
1971 "Robustness in regression analysis." Pp. 118-46 in Herbert L. Costner (ed.), *Sociological Methodology 1971*. San Francisco: Jossey-Bass.
- Boyle, Richard P.
1970 "Path analysis and ordinal data." *American Journal of Sociology* 75:461-80.
- Centers, Richard
1949 *The Psychology of Social Classes: A Study of Class Consciousness*. New York: Russell and Russell.
- Chinoy, Ely
1955 *Automobile Workers and the American Dream*. Garden City: Doubleday.
- Commager, Henry Steele
1974 "English traits: one hundred years later." Pp. 749-59 in Henry Steele Commager (ed.), *Britain Through American Eyes*. New York: McGraw-Hill.
- Cook, Karen S.
1975 "Expectations, evaluations and equity." *American Sociological Review* 40:372-88.
- Cronbach, Lee J.
1951 "Coefficient alpha and the internal structure of tests." *Psychometrika* 16:297-334.
- Feagin, Joe R.
1972 "Poverty: we still believe that God helps those who help themselves." *Psychology Today* 6:101-10, 129.
- Form, William H., and Joan Rytina
1969 "Ideological beliefs on the distribution of power in the United States." *American Sociological Review* 34:19-31.

- Gans, Herbert J.
1973 *More Equality*. New York: Vintage.
- Glyn, Anthony
1970 *The British: Portrait of a People*. New York: Putnam's.
- Goldthorpe, John H., and David Lockwood
1963 "Affluence and the British class structure." *Sociological Review* 11:133-63.
- Goldthorpe, John H., David Lockwood, Frank Bechhofer and Jennifer Platt
1969 *The Affluent Worker in the Class Structure*. Cambridge, Ma.: Cambridge University Press.
- Hamilton, Richard F.
1972 *Class and Politics in the United States*. New York: Wiley.
- Harrington, Michael
1962 *The Other America*. New York: Macmillan.
- Heise, David R., and George W. Bohrnstedt
1970 "Validity, invalidity, and reliability." Pp. 104-29 in Edgar F. Borgatta and George W. Bohrnstedt (eds.), *Sociological Methodology* 1970. San Francisco: Jossey-Bass.
- Huber, Joan, and William H. Form
1973 *Income and Ideology: An Analysis of the American Political Formula*. New York: Free Press.
- Hyman, Herbert H.
1953 "The value systems of different classes." Pp. 426-42 in Reinhard Bendix and Seymour Martin Lipset (eds.), *Class, Status, and Power*. Glencoe: Free Press.
- Jasso, Guillermina, and Peter H. Rossi
1977 "Distributive justice and earned income." *American Sociological Review* 42:639-51.
- Kerckhoff, Alan C.
1977 "The realism of educational ambitions in England and the United States." *American Sociological Review* 42:563-71.
- Kristol, Irving
1968 "Equality as an ideal." Pp. 108-11 in David L. Sills (ed.), *International Encyclopedia of the Social Sciences*, Vol. 5. New York: Macmillan.
- Labovitz, Sanford
1967 "Some observations on measurement and statistics." *Social Forces* 46:151-60.
1970 "The assignment of numbers to rank order categories." *American Sociological Review* 35:515-24.
1971 "In defense of assigning numbers to ranks." *American Sociological Review* 36:521-2.
- Lane, Robert E.
1959 "The fear of equality." *American Political Science Review* 53:35-51.
- Lipset, Seymour Martin
1963 "The value patterns of democracy: a case study in comparative analysis." *American Sociological Review* 28:515-31.
1966 "Value systems, class, and the democratic polity: the United States and Great Britain." Pp. 161-71 in Reinhard Bendix and Seymour Martin Lipset (eds.), *Class, Status, and Power*. 2nd ed. New York: Free Press.
- 1967 *The First New Nation*. Garden City: Doubleday.
- 1968 "Anglo-American society." Pp. 289-302 in David L. Sills (ed.), *International Encyclopedia of the Social Sciences*, Vol. 1. New York: Macmillan.
- McClosky, Herbert
1964 "Consensus and ideology in American politics." *American Political Science Review* 58:361-82.
- MacDonald, Dwight
1963 "Our invisible poor." *New Yorker* (January 19):82-92.
- McKenzie, Robert, and Allan Silver
1968 *Angels in Marble: Working Class Conservatives in Urban England*. Chicago: University of Chicago Press.
- Maillaud, Pierre
1946 *The English Way*. London: Oxford University Press.
- Mau, James A., Richard J. Hill and Wendell Bell
1961 "Scale analyses of status perception and status attitude in Jamaica and the United States." *Pacific Sociological Review* 4:33-40.
- Meier, Dorothy Louise
1962 *Anomia, Life Chances, Perceived Achievement, and Modes of Adaptation*. Ph.D. dissertation, Department of Anthropology and Sociology, University of California, Los Angeles.
- Miller, S. M., and Martin Rein
1966 "Poverty, inequality and policy." Pp. 426-516 in Howard S. Becker (ed.), *Social Problems*. New York: Wiley.
- Moskos, Charles C., Jr., and Wendell Bell
1967 "Attitudes toward equality." Pp. 100-14 in Wendell Bell (ed.), *The Democratic Revolution in the West Indies*. Cambridge, Ma.: Schenkman.
- Nordlinger, Eric A.
1967 *The Working Class Tories: Authority, Defiance and Stable Democracy*. London: MacGibbon and Kee.
- Nowak, Stefan
1969 "Changes and social structure in social consciousness." Pp. 235-47 in Celia S. Heller (ed.), *Structured Social Inequality: A Reader in Comparative Social Stratification*. London: Collier-Macmillan.
- Petchen, Martin
1961 "A conceptual framework and some empirical data regarding comparisons of social rewards." *Sociometry* 24:136-56.
- Rainwater, Lee
1974 *What Money Buys: Inequality and the Social Meaning of Income*. New York: Basic Books.
- Rawls, John
1971 *A Theory of Justice*. Cambridge, Ma.: Belknap.
- Robinson, Robert V., and Jonathan Kelley
1977 "Marx and Dahrendorf on income inequality, class consciousness and class conflict: an empirical test." Paper read at the annual

- meeting of the American Sociological Association, Chicago.
- Rokeach, Milton
1973 *The Nature of Human Values*. New York: Free Press.
- Rose, Richard
1964 *Politics in England: An Interpretation*. Boston: Little, Brown.
1974 *Politics in England: An Interpretation*. 2nd ed. Boston: Little, Brown.
- Rose, Richard, and Dennis Kavanagh
1976 "The monarchy in contemporary political culture." *Comparative Politics* 18:548-76.
- Roshwald, Mordecai
1973 "Order and over-organization in America." *British Journal of Sociology* 24:30-42.
- Runciman, W. G.
1966 *Relative Deprivation and Social Justice: A Study of Attitudes to Social Inequality in Twentieth-Century England*. Berkeley: University of California Press.
- Seeman, Melvin
1960 *Social Status and Leadership*. Columbus: Bureau of Educational Research and Service, Ohio State University.
- Treiman, Donald J.
1977 *Occupational Prestige in Comparative Perspective*. New York: Academic.
- Treiman, Donald J., and Kermit Terrell
1975 "The process of status attainment in the United States and Great Britain." *American Journal of Sociology* 81:563-83.
- Turner, Ralph H.
1958 "Life situation and subculture: a comparison of merited prestige judgments by three occupational classes in Britain." *British Journal of Sociology* 9:299-320.
1960 "Sponsored and contest mobility and the school system." *American Sociological Review* 25:855-67.
- Useem, Ruth Hill, and John Useem
1954 "Images of the United States and Britain held by foreign-educated Indians." *Annals of the American Academy of Political and Social Science* 295:73-82.
- Weick, Karl E.
1966 "The concept of equity in the perception of pay." *Administrative Science Quarterly* 11:414-39.
- Westergaard, John, and Henrietta Resler
1975 *Class in a Capitalist Society: A Study of Contemporary Britain*. London: Heinemann.
- Williams, Robin
1951 *American Society: A Sociological Interpretation*. New York: Knopf.

MANUSCRIPTS FOR THE ASA ROSE SOCIOLOGY SERIES

Manuscripts (100 to 300 typed pages) are solicited for publication in the *ASA Arnold and Caroline Rose Monograph Series*. The Series welcomes a variety of types of sociological work—qualitative or quantitative empirical studies, and theoretical or methodological treatises. An author should submit three copies of a manuscript for consideration to the Series Editor, Professor Robin M. Williams, Jr., Department of Sociology, Cornell University, Ithaca, New York 14853.

INFORMED CONSENT: CONSEQUENCES FOR RESPONSE RATE AND RESPONSE QUALITY IN SOCIAL SURVEYS*

ELEANOR SINGER
Columbia University

American Sociological Review 1978, Vol. 43 (April):144-162

The study described in this paper was designed to measure the effects of variations in (1) the amount of information provided to respondents about the content of a survey ahead of time, (2) the assurance of confidentiality given to respondents, and (3) the request for and timing of a signature to document consent on three aspects of social surveys: (a) overall response rate, (b) response rates to individual questions, and (c) response quality. That is, the study was designed to measure the impact of informed consent procedures on response rates and response quality in social survey research. For this purpose, a fully crossed 2 x 3 x 3 factorial design was imposed on a national probability sample of 2,384 potential respondents. The questionnaire consisted of a large number of items in such presumably sensitive areas as drinking, marijuana use, sexual behavior, and mental health, in addition to more conventional questions about leisure activities. Major findings can be summarized as follows: 1. The overall response rate to the survey was 67%; of the three variables investigated, only the request for a signature had a significant effect on the probability of responding. Seventy-one percent of those not asked for a signature were interviewed, compared with 64% and 65% of those asked to sign before and those asked to sign afterwards. 2. Only the assurance of confidentiality had a significant effect on item nonresponse. Despite the sensitive nature of the interview, nonresponse to individual questions was very low. On those questions to which the nonresponse rate totaled more than 3%—all of them questions about behavior rather than attitudes—respondents given an assurance of absolute confidentiality had a lower nonresponse rate than those in two other experimental groups, in some cases by a statistically significant margin. 3. None of the three independent variables had either consistent or large effects on the quality of response. However, there are suggestions in the data that asking for a signature before the interview has a sensitization effect, so that better data are obtained if the respondent is asked to sign a consent form afterwards.

Our age, as Bernard Barber (1973) put it, is the "age of civil rights"; it is an age of increasing concern for the protection of individual rights against institutional encroachment and abuse. One area of concern is that of research involving human subjects, primarily but not exclu-

sively biomedical research; and one issue that has gained prominence within that area is the issue of informed consent. That issue has been sharpened by existing and proposed regulations designed to protect the rights of human subjects.

Among scientists, debate over these regulations has been conducted, largely without benefit of empirical evidence, between those who champion the new rules on ethical grounds (e.g., Barber et al., 1973: Chap. 3; Gray, 1975: Chaps. 8 and 9; Warwick, 1975:38ff.) and those who oppose them in the belief that they will destroy the possibility of doing needed research. The present study, funded by a grant from the National Science Foundation and carried out under the auspices of the National Opinion Research Center, was designed to provide evidence bearing on this issue, so far as social survey research is concerned.

Specifically, the study was designed to

* The work reported here owes much to many people. I would like especially to thank James R. Murray and Martin R. Frankel for helping to design the experiment; Steven M. Cohen for advice in planning the analysis; Luane Kohnke-Aguirre for carrying it out; and Norman M. Bradburn, Charles D. Cowan, Kenneth Prewitt, Arthur H. Stinchcombe, and Seymour Sudman, for reading and commenting on an earlier draft. I would also like to express my appreciation for the grant from the National Science Foundation (No. Soc 75-22889) which made it possible to translate a proposal for research into a completed experiment. The study was carried out as a Visiting Senior Study Director at the National Opinion Research Center in New York.

A preliminary report of these findings was given at the American Sociological Association Meeting, Chicago, 1977.

investigate the effects of three factors that, together, may be said to constitute informed consent procedures in face-to-face interviews with a random sample of the adult population.

The first of these factors was the amount of information given to respondents ahead of time about the content of the interview. Conventional survey wisdom advocates keeping the introduction short, so as not to lose the respondent's interest or attention; and some evidence from experiments with mail questionnaires suggests that a general explanation of purpose is preferable to a more detailed one, which may antagonize some respondents (Blumberg et al., 1974). The advice in the most recent edition of the *Handbook of Social Psychology*, as in earlier editions, is to keep subjects in laboratory experiments ignorant of the true purposes of the research (Aronson and Carlsmith, 1969:61).

At the same time, some investigators support fuller disclosure of research purposes to respondents. Jourard (1964; 1968; Jourard and Friedman, 1970), for example, has argued that the most powerful determinant of self-disclosure by experimental subjects is self-disclosure on the part of the investigator. Although his methods do not appear to be easily adapted to the usual laboratory situation, they may have more direct implications for survey research. If Jourard is correct, then full disclosure of the purpose of a survey—though not necessarily of specific hypotheses—actually may result in a higher response rate and less response distortion than attempts at deception. Some support for this derives from a study of telephone household screening (Hauck and Cox, 1974), in which refusals were reduced after respondents had been given a more nearly complete and accurate description of the study's purpose.

Accordingly, we constructed two descriptions of the content of the interview, to be read to respondents ahead of time. Half the respondents were given a brief, vague description of the survey as a study of leisure time and the way people are feeling. The other half were given a fuller description of the interview, which contained a large number of questions gener-

ally considered sensitive. These respondents were told:

We're conducting a national survey about how people are feeling in general and about the kinds of activities people do in their leisure time—that is, their spare time when they are not working. There are questions about your moods, and about the time you spend watching television or going to sports events, about your social activities, and some about your use of alcoholic drinks. We also ask a few questions about sex.

Aside from being shorter, the short introduction is, essentially, a deceit condition. Though the deceit is mild, the information given to the respondent ahead of time is not consistent with the relatively heavy emphasis on drinking, sex, and mental health in the interview.

We expected that disclosing the sensitive content of the questions would increase their threat for some people but engender an atmosphere of trust for others. In addition, we anticipated that the more detailed introduction would make the attitudes and behaviors asked about more salient, thereby increasing the accuracy with which they were reported. Taking all these cross-pressures into account, we expected no significant differences in initial refusal between information conditions, but more accurate reporting (i.e., more admission of undesirable behavior) once the respondent had agreed to participate.

The second factor that was varied experimentally in the study was the assurance of confidentiality given to respondents. It has become increasingly clear that although some research organizations, such as NORC, routinely promise to protect the confidentiality of respondent replies, such guarantees ordinarily have no legal standing; the relation between researcher and respondent is not recognized as privileged.¹ If records are

¹ Current federal law provides examples of several types of protections of research data from compelled disclosure—e.g., the statute prohibiting disclosure of Census records (13 U.S.C. 8, 9); the limited protection enjoyed by the National Center for Health Statistics for all individually identifiable research information it collects (42 U.S.C. 242m). Patient records maintained in connection with any drug abuse program or research activity conducted, regulated, or directly or indirectly assisted by a federal agency

subpoenaed, there is ultimately nothing, short of going to jail, that the researcher can do to redeem the promise of confidentiality made to respondents. If there is no need to identify respondents for administrative purposes or follow-up studies, the problem of confidentiality can sometimes be handled by destroying overtly identifying information. But in small and specialized populations, even this strategy will not prevent the identification of the respondent from other bits of information routinely punched on IBM cards.²

Unlike the argument for other aspects of informed consent, which is advanced on ethical grounds only, that for confidentiality is advanced on pragmatic grounds as well. That is, not only is breach of confidentiality a risk to which respondents are exposed by virtue of participating in the survey; but, it has been argued, the respondent will not give valid information without the promise that the confidentiality of his or her replies will be maintained.

Although no prior research on the effects of confidentiality existed when this study was designed, research on the closely related factor of anonymity suggests that response rates are not much affected either by the presence or absence of an identifying code, or by the form of identification used (Blumberg et al., 1974; Erdos and Regier, 1977; Mason et al., 1961; Mitchell, 1939).

Aside from its effect on response rates, lack of anonymity has frequently been considered a potential source of bias in mail surveys. Fuller (1974), however, concluded from her review of the literature that the risk of significant bias was relatively small. While a few researchers

have found some systematic bias, others have not, even when the information was somewhat sensitive (see Ash and Abramson, 1952; Becker and Bakal, 1970; Fischer, 1946; Fuller 1974; Haberman et al., 1972; Hamel and Reif, 1952; King, 1970; Kosen et al., 1970; Rosen, 1960; Wildman, 1977).

In order to investigate the effects of variations in confidentiality, one-third of the respondents in the present study were told nothing at all about the confidentiality of their replies; one-third were given an absolute assurance of confidentiality ("Of course, your answers will remain completely confidential"); and one-third were given a qualified assurance of confidentiality ("Of course, we will do our best to protect the confidentiality of your answers, except as required by law").

The final factor that was varied in the study was whether or not a signature was required to document consent, and if so, whether the request for a signature came before or after the interview.

It is generally assumed that requiring a signature to document consent will lower response rate, though the research on anonymity, which also bears indirectly on this problem, suggests that this effect may well be small. Deferring a signature until after the interview has been completed has several potential advantages: (1) in some sense, truly informed consent can only be given after the respondent has heard the actual questions; (2) responses, as distinct from response rate, will be unaffected by a request for a signature which is made at the end of the interview; (3) response rate may be protected to some extent if the respondent, having invested time in giving the interview, is reluctant to see it wasted by refusing to sign the consent form.

Aside from these three factors, which can also be thought of as representing different levels of risk or cost to respondents, certain elements of the introduction were kept constant. All respondents were told that the study was being done by the National Opinion Research Center; that the interview would take about half an hour, and that they would then be asked to fill out a short self-administered form; that participation was voluntary and that they

or department enjoy limited protection (42 U.S.C. 4582; 21 U.S.C. 1175). The Secretary of DHEW may authorize researchers engaged in mental health or alcohol or drug abuse research to withhold names or identifying characteristics of subjects, and this immunity covers them in any federal, state, or local civil, criminal, administrative, legislative, or other proceeding (42 U.S.C. 4582). But these protections are rarely absolute, and they cover only limited types of data.

² For a review of these and other problems related to confidentiality, see Committee on Federal Agency Evaluation Research (1975), Nejeleski (1976), and Privacy Protection Study Commission (1977: Chap. 15).

Table 1. Research Design for Measuring the Effects of Variations in Amount of Information, Assurance of Confidentiality, and Request for a Signature

	Detailed, Long Information			Vague, Short Information		
	Absolute Confidentiality	Qualified Confidentiality	No Mention of Confidentiality	Absolute Confidentiality	Qualified Confidentiality	No Mention of Confidentiality
Signature before	115*	115	115	115	115	116
Signature after	115	115	116	115	116	115
No signature	115	116	116	116	115	116

*The figures shown refer to the number of cases assigned to each condition.

could refuse questions within the interview. Every introduction also included a plea for honesty of response if the person decided to participate.

The three factors described above were combined in a 2 x 3 x 3 factorial design, yielding 18 different introductions to respondents (see Table 1). These introductions were stapled to the household enumeration folders assigned to each interviewer, along with instructions for answering questions and objections, also specifically tailored to each version of the introduction. Interviewers were instructed to read the introduction in its entirety to the selected respondent, and to give standardized replies to questions that might be raised about content, confidentiality, or signature.

A national probability sample of 2,084 (drawn within 50 primary sampling units of NORC's master sample) was used for the study.³ Interviewers were required to list all household members and to select the appropriate respondent according to a sampling table. Each sample line (household) had been assigned a randomly

selected experimental treatment in the central office. Every interviewer was assigned 31 sample lines, and every interviewer's assignment included all 18 versions of the introduction.⁴

The interview schedule was substantially the same as that used a year earlier by Sudman and Bradburn in a study of the effects of question threat and question wording.⁵ Following some questions about conventional leisure activities were sections dealing with emotional well-being and mental health, drinking, marijuana use, and sexual behavior. A number of demographic items, among them income, were also included. The questionnaire was designed to permit investigation of the ef-

³ The universe is defined as the noninstitutionalized population 18 and over resident in the 48 contiguous states. For a more detailed description of the sample, see National Data Program for the Social Sciences (1976:94-5).

In order to detect a difference of 7% or more between extreme levels of a main effect at an alpha level of 5% under the worst possible assumption—that is, with proportions symmetrical around 50%—we calculated that a completed sample of 1,200 would be necessary. The gross sample size was determined by assuming that 10% of the assigned lines would be lost because of vacancies and other changes in dwelling units, and that because of the sensitive nature of the questionnaire and some limitations on follow-up procedures, the completion rate on the remaining cases would average 65%. In fact, the completion rate was 67%.

⁴ Interviewer effects were balanced in the design by having each interviewer administer all 18 experimental treatments, with order of administration randomized. A total of 2,077 lines was selected in 67 segments (clusters) of 50 primary sampling units. Within each segment, 31 sample lines (households) were assigned to one interviewer, and each of these households was randomly assigned a predesignated experimental treatment. For example, in the first segment, the first 18 lines selected received treatments 1 through 18, randomized to balance order effects. The next 13 lines were assigned a subset of 13 treatments randomly drawn from the complete set of 18. In the next cluster, the remaining 5 treatments were assigned together with a complete set of 18 treatments, plus a randomly drawn set of 8 treatments, for a total of 31 lines. Whenever a subset n of the 18 treatments was assigned to a cluster, leaving a pool of $18-n$ treatments, the remaining $18-n$ treatments were assigned to the next cluster. Interviewers were instructed to make initial contact with the assigned households in their segment beginning with the lowest case number and ending with the highest, although we recognized that some deviations from the assigned order were likely to occur.

⁵ National Science Foundation Grant No. GS-43245, Seymour Sudman and Norman M. Bradburn, principal investigators. For some of the findings from this study, see Blair et al. (1977), Bradburn et al. (1978), and Sudman et al. (1977).

Table 2. Questions Potentially Affected by Informed Consent Procedures

Attitudes, Opinions		Behavior, Fact	
General	Survey-Specific	General	Survey-Specific
1	2	3	4
<i>Sensitive:</i>			
E.g., questions about satisfaction with finances	E.g., questions about emotional state	E.g., questions about income	E.g., questions about alcohol consumption
5	6	7	8
<i>Nonsensitive:</i>			
E.g., questions about party preference	E.g., preferred leisure activities	E.g., questions about education	E.g., amount of time watching TV

fect of informed consent procedures on eight types of questions: sensitive and nonsensitive survey-specific and general questions about attitudes and about behavior (see Table 2).

Also measured in the study were respondents' reactions to the interview, ascertained by means of a self-administered questionnaire filled out immediately following the interview and handed to the interviewer in a sealed envelope, and interviewers' expectations about and reactions to the study, assessed just prior to and after the completion of field work. After all interviewing had been completed, each respondent was sent a letter thanking him or her for participating, explaining that the study had had a methodological as well as a substantive purpose, and briefly describing that purpose.⁶

With one exception, only experienced NORC interviewers—almost a third of whom had worked on the Sudman-Bradburn study a year earlier—were used on this survey. They were trained in the special experimental procedures for the study through a combination of written materials, group telephone briefings by area supervisors, and specially developed training exercises which had to be completed before interviewing could begin. Interviews were edited in the New York office, and 20% of each interviewer's cases were validated.

Interviewers were told about the

methodological purposes of the study but not about any specific hypotheses, and were urged to keep an open mind about the effects of the experimental variables. Those who felt seriously uncomfortable with either the substantive or the methodological aspects of the study were asked not to take on this particular assignment; about five withdrew for this reason.

Several facts give us confidence that interviewers followed the procedures specified. First, the two variations in the description of content had been tried earlier, in the Sudman-Bradburn study, with essentially the same results as those reported here. Second, evidence that the respondent was asked to sign the consent form is provided by the forms themselves, unless we assume that interviewers forged some signatures.⁷ No such evidence is available for confidentiality. However, we assume that distortions with respect to this variable, if they occurred, would have been in the direction of promising confidentiality to those respondents who should not have received such an assurance. Therefore, we will be on safe grounds if we treat the findings on confidentiality as minimal effects.

This article discusses the effect of each of the three experimental variables on three different types of outcomes: overall response rate to the survey, response rates to individual questions, and response quality. Discussion of respondent

⁶ This method was decided on in preference to personal debriefing in order to keep interviewers from receiving feedback that might affect their future performance.

⁷ The only protection against this lies in training and supervision, including validation of each interviewer's work.

Table 3. Reasons Given by Interviewers for Failing to List Household Members

Reason	Percent of Sample	(N)
Vacant dwelling unit	4.9%	(103)
No dwelling unit	6.0	(126)
Refused	8.6	(179)
(Breakoff)		(1)
Not home	1.5	(32)
No English	0.9	(18)
Unavailable	0.3	(7)
Other	0.1	(3)
Listed	77.5	(1,615)
Total	99.8	(2,084)

reactions and interviewer expectations is reserved for future articles.

Effects on Response Rate to the Questionnaire

The overall response rate to the questionnaire was 67%; 87% of all households in the gross sample,⁸ minus only those which were vacant or no-dwelling units, were successfully listed; 77% of those listed were actually interviewed.

Distribution of the assigned sample among the 18 different experimental conditions is shown in Table 1. Chi-squares computed on the distribution of the actual gross sample and on the distribution of all those whose households were successfully listed indicate that they do not depart significantly from a distribution assuming an equal number of cases in each cell. That is, there were no differential losses from the sample during listing. Reasons for failing to list household members or to complete interviews with those listed are summarized in Tables 3 and 4.

Of the three experimental variables, only the request for a signature had a statistically significant effect on the probability of responding, as evaluated by analysis of variance.⁹ Seventy-one per-

⁸ The gross sample consists of the number of assigned lines (N = 2,077), plus previously unlisted dwelling units (N = 38), minus 31 lines lost when one segment was dropped without replacement because of a very large proportion of non-English-speaking residents.

⁹ The statistical tests used throughout assume simple random sampling, although the sample is stratified and clustered. For practical purposes, this yields a conservative estimate of the effect of in-

Table 4. Reasons for Failure to Complete Main Interview

Reason	Percent of Listed Households	(N)
Refused*	15.6%	(253)
Breakoff	0.2	(4)
No English	1.2	(19)
Not home	1.2	(19)
Too ill	1.7	(27)
Unavailable	0.4	(7)
Invalid interview	1.0	(16)
Wrong respondent	1.0	(17)
Ineligible for sample	0.4	(6)
Interviewed	77.2	(1,247)
Total	99.9	(1,615)

* Includes 74 respondents who were interviewed but refused to sign the consent form.

cent of those not asked to sign were interviewed, compared with only 64% of those asked before, and 65% of those asked after (see Table 5). Even so, it should be noted that the refusal was limited to the signature itself. Only a handful of respondents actually refused to be interviewed; the rest agreed to the interview but refused to sign the consent form, or signed after the interview rather than before.¹⁰ None of the two- or three-way interactions among independent variables was significant. Because those respondents who refused to sign the consent form were actually interviewed, it is possible to examine the effects of content and confidentiality alone—i.e., ignoring the effect of the signature variable altogether—by including those who refused to sign the consent form in the analysis. The overall response rate, under these circumstances, is of course higher (71% vs. 67%) but neither content nor confidentiality has an effect on it.

Since the data fit an additive model, it is also possible to estimate the cumulative effect of all three informed consent variables by summing the adjusted deviations associated with the short content, absolute confidentiality, and no-signature

formed consent procedures, in the sense that there are likely to be fewer rather than more true effects than those reported here.

¹⁰ If those who refused to sign the consent form but agreed to be interviewed are included, the response rate for the signature-before condition is 71%, and for the signature-after condition, 70%.

Table 5. The Effect of Variations in Information about Content, Assurance of Confidentiality, and Request for a Signature on the Probability of Responding to the Interview

Independent Variable	Probability of Response ^a	(N) ^b	Eta
Content			
Long	.68	(922)	
Short	.66	(933)	.03
Confidentiality			
No mention	.68	(634)	
Qualified	.66	(603)	
Absolute	.67	(618)	.01
Signature			
Not asked	.71	(625)	
Asked before	.64	(608)	
Asked after	.65	(622)	.07*

^a Adjustment (by multiple classification analysis) for the effects of the other two independent variables does not affect the results to two decimals.

^b The response rate has been computed with the gross sample (N=2,084) minus households which were vacant or no dwelling units (N=229), or 1,855 households, as the denominator.

* $p(F) = .01$.

conditions—in other words, the standard survey introduction—and contrasting them with the summed deviations for the long, qualified, and signature-before conditions. The estimated response rates are 70% and 64%, respectively. That is, the estimated response-rate cost associated with all three elements of informed consent is no greater than that associated with the request for a signature alone.¹¹

It can be argued that there is no way for the experimental variables to affect the response rate unless the interviewer has actually read the appropriate introduction.¹² Accordingly, we asked interviewers to indicate how much of the introduction had been read to each person who

¹¹ Although the comparison is hazardous for several reasons, it might be noted that the comparably defined response rate to that part of the General Social Survey based on a probability sample, carried out by NORC only a few months earlier, was 71%.

¹² This is not necessarily true. If the interviewers had misgivings about certain introductions, they might have been less persistent in attempting to locate a respondent or less effective in obtaining an interview. Or, respondents might have asked questions about some of the experimental variables and declined to be interviewed on the basis of the responses they were given.

refused to be interviewed. Overwhelmingly (65% of the time) they reported that the refusal occurred before any part of the introduction had been read. The implications of this are considered later. Nevertheless, it was possible to examine refusals by experimental condition among those respondents who had heard the relevant portion of the introduction. For example, in order to examine the effect of information about content on refusals, we included all those to whom the interviewer had read at least that section of the introduction. To examine the effect of confidentiality, we included all those who had heard that portion of the introduction, and so on. We assumed that, as specified, the entire introduction had been read to all respondents who had been interviewed.

Table 6 shows the relation between experimental treatment and refusal, for those to whom the relevant portions of the introduction had been read. That table confirms the conclusion drawn on the basis of the earlier analysis: namely, that only the request for a signature has a discernible effect on overall response.

Effects on Response Rates to Individual Questions

Although the interview schedule included many questions ordinarily considered sensitive or threatening, the rate of

Table 6. Refusal by Experimental Condition, among Those Who Heard Relevant Portion of Introduction

Experimental Variable	Percent Who Refused	Chi-Square
Content		
Long	5% (693) ^a	
Short	4 (688)	1.68; n.s.
Confidentiality		
No mention	4 (475)	
Qualified	3 (443)	
Absolute	3 (444)	0.61; n.s.
Signature		
Not asked	0 (446)	
Asked before	11 (443)	
Asked after	7 (439)	48.49; $p < .01$

^a Numbers shown include only potential respondents who heard the relevant portion of the introduction.

nonresponse to individual items was very low. Only two questions elicited a total nonresponse rate—i.e., no answer, don't know, and not asked¹³—of 10% or more: a question about income (11% nonresponse), and one about masturbation (10% nonresponse). Most items had nonresponse rates below 2%, and we did not examine the effect of the experimental variables on these.

We did, however, examine all questions with a total nonresponse rate of more than 3%.¹⁴ All of these, it should be noted, asked about some item of behavior; none of the attitude questions elicited a nonresponse rate as high as this.

The results of this analysis are quickly summarized (for details, see Tables 7, 8, and 9). The amount of information given respondents ahead of time has no statistically significant or consistent effects on nonresponse to individual questions; in only one of twelve comparisons do differences in nonresponse exceed three percentage points.

The assurance of confidentiality, on the other hand, does appear to affect the rate of nonresponse to individual questions. With one exception, respondents who were told that their answers would remain completely confidential had the lowest nonresponse rate of any of the three groups. In five of the twelve comparisons, differences among the three groups were statistically significant.¹⁵

¹³ Because of the small number of responses coded as refused, don't know, and not asked, all three have been included in the category of nonresponse. Conclusions do not change if the "not asked" category is excluded from analysis, and there is some justification for including it as a measure of effects on interviewers. Such effects can, of course, affect response rate, even though indirectly.

¹⁴ Nonresponse rates to eight of the twelve items are from two to five percentage points higher in the present study than in the Sudman-Bradburn study referred to earlier. This result may be due in part to the fact that, with a few exceptions, we used that version of the Sudman-Bradburn questionnaire associated with the lowest estimates of sensitive behavior (see Blair et al., 1977). In part, however, it may also reflect the cumulative cost of the experimental variables used in the present study.

¹⁵ No clear ordering among the three experimental groups can be specified as an alternative to the hypothesis of no difference. Prior research and theory suggest that respondents given an absolute assurance of confidentiality may be more willing to answer

The final variable whose effect on item nonresponse we examined was the request for a signature and its timing. Since the present survey was an experiment rather than a study which actually required informed consent, we conducted interviews with those who refused to sign and coded their refusal. We are therefore able to compare the responses given to individual questions by those who agreed to sign and by those who refused.

As it turns out, the biggest differences in nonresponse are between those who refused to sign the consent form and those who either signed or were not asked to sign. Differences in nonresponse between the latter groups are very small—generally, two percentage points or less. Those who were asked to sign before the interview and agreed to do so appear to be somewhat less likely to refuse to answer questions thereafter. At the same time, those asked for a signature beforehand were somewhat more likely to refuse to sign at all: 10% of those asked to sign the consent form before the interview refused, compared with 7% of those asked to sign afterwards ($z = 1.68$; $p < .10$).¹⁶

An alternative method of analysis involves summing the number of nonresponses to all four items asked of the total sample, and then comparing this index of nonresponse across experimental conditions. (Since the remaining questions are contingent on the others, they cannot be included in the index.) The data are shown in Table 10, and lead to the same conclusions as do those reported above.

The high item nonresponse rate of those who had refused to sign the consent form led us to ask whether refusers are a distinct group, inclined to say "no" to questions or requests which they construe as

questions than those who are not, without indicating whether those to whom the matter is not mentioned would be more willing to answer questions than those to whom only a qualified assurance is given, or vice versa.

¹⁶ Among those who did sign the consent form, there were no differences in reluctance, as noted by interviewers at the conclusion of the interview. Eighty-seven percent of those who signed after the interview were judged not to have been reluctant at all; 12% were somewhat reluctant; and 1%, very reluctant. Among those who signed before, the corresponding percentages were 88, 11, and 1.

Table 7. Nonresponse ^a to Sensitive ^b Questions, by Information Condition

Question	Percent Nonresponse		Significance ^c
	Long Introduction	Short Introduction	
1. Ever smoked marijuana 3 times a week or more? ^d	4.7% (149)	9.3% (161)	n.s.
2. Number of pipes, joints smoked per time? ^d	4.7 (149)	6.2 (161)	n.s.
3. Engaged in petting or kissing in past month?	7.3 (658)	5.3 (663)	n.s.
4. How often? ^d	7.5 (465)	8.8 (445)	n.s.
5. Petted, kissed in past 24 hours? ^d	2.8 (465)	4.7 (445)	n.s.
6. Intercourse in past month?	9.9 (658)	7.8 (663)	n.s.
7. How often? ^d	11.3 (407)	8.7 (393)	n.s.
8. Intercourse in past 24 hours? ^d	6.1 (407)	3.3 (393)	n.s.
9. Masturbated in past month?	12.5 (658)	9.7 (663)	n.s.
10. How often? ^d	7.0 (57)	8.3 (36)	n.s.
11. Masturbated in past 24 hours? ^d	5.3 (57)	5.6 (36)	n.s.
12. Earned income in past year?	10.5 (658)	10.1 (663)	n.s.

^a "Nonresponse" includes refusal, don't know, and not asked.

^b All questions on which more than 3% of responses were coded as missing data are included in this table. With one exception, they come from the "sensitive, survey-specific, factual" question category. The exception is the question on income, which is general rather than survey-specific.

^c Significance levels are based on chi-square.

^d For these questions we give the conditional probability of refusing to answer, among those who answered the filter question affirmatively.

^e On this particular question, 16 of the 17 missing responses are attributable to interviewer omissions rather than respondent refusals.

sensitive, or whether, perhaps, asking respondents to sign a consent form ahead of time had sensitized them to the content of the interview, so that they were more likely to refuse to answer particular questions during the interview itself. If the first hypothesis were true, there should be little, if any, difference in nonresponse to specific questions between those who refused to sign before, and those who refused to sign afterwards. If the second hypothesis were true, high item nonresponse rates should be characteristic only of those who refused to sign a consent form before the interview.

The relevant data are shown in Table 11. Although the numbers on which the percentages in each cell are based are now very small, the figures support the interpretation that refusers are a distinct group. Eight of the twelve comparisons indicate that there are no differences in response tendencies to specific items between those who refused to sign before and those who refused to sign afterwards. The four that do indicate such a

difference invariably produce a larger nonresponse tendency for the group refusing to sign the consent form after the interview. In other words, rather than the request for a signature sensitizing people to the content of the interview, it appears to be true that those who reacted negatively to some of the questions, subsequently refused to sign the consent form. For the most part, however, signing the consent form appears to function simply as another sensitive question, so that those who refuse the questions, refuse to sign, and vice versa.

Age, sex, race, and number of years of school were examined as potential predictors of refusal to sign the consent form. Table 12 indicates that men and women were equally likely to refuse, and that differences between black and white respondents were insignificantly small. The relation between education and refusal may be curvilinear, with those respondents having less than a high school education, and those with a college degree or more, more likely to refuse than intermediate

Table 8. Nonresponse ^a to Sensitive ^b Questions, by Confidentiality Condition

Question	Percent Nonresponse				Significance Level ^c
	No Mention of Confidentiality	Qualified Confidentiality	Absolute Confidentiality		
1. Ever smoked marijuana 3 times a week or more? ^d	7.5% (120)	10.6% (94)	3.1% (96)		n.s.
2. Number of pipes, joints smoked per time? ^a	5.0 (120)	9.6 (94)	2.1 (96)		n.s.
3. Engaged in petting or kissing in past month?	8.5 (457)	6.0 (430)	4.1 (434)		<.05
4. How often? ^d	7.0 (315)	10.0 (285)	7.5 (306)		n.s.
5. Petted, kissed in past 24 hours? ^d	4.4 (315)	4.2 (285)	2.6 (306)		n.s.
6. Intercourse in past month?	10.7 (457)	9.5 (430)	6.2 (434)		=.05
7. How often? ^d	10.3 (271)	11.7 (257)	8.1 (272)		n.s.
8. Intercourse in past 24 hours? ^d	6.3 (271)	5.8 (257)	2.2 (272)		=.05
9. Masturbated in past month?	14.2 (457)	10.7 (430)	8.1 (434)		<.05
10. How often? ^d	6.3 (32)	11.5 (26)	5.7 (35)		n.s.
11. Masturbated in past 24 hours? ^d	6.3 (32)	7.7 (26)	2.9 (35)		n.s.
12. Earned income in past year?	12.9 (457)	10.7 (430)	7.1 (434)		<.05

^a "Nonresponse" includes refusal, don't know, and not asked.

^b All questions on which more than 3% of responses were coded as missing data are included in this table. With one exception, they come from the "sensitive, survey-specific, factual" question category. The exception is the question on income, which is general rather than survey-specific.

^c Significance levels are based on chi-square.

^d For these questions we give the conditional probability of refusing to answer, among those who answered the filter question affirmatively.

^e On this particular question, 16 of the 17 missing responses are attributable to interviewer omissions rather than respondent refusals.

groups, the highest refusal rate occurring among those with the least education.

The strongest relationship, however, is between refusal and age. With one reversal, as age increases, so does refusal to sign the consent form, and when all four variables were entered simultaneously into a regression equation, only age retained a significant effect. However, all four together account for less than 1% of the variance in refusal.

Effects on Quality of Response

Previous research has indicated that for sensitive or threatening questions, more reporting is better reporting.¹⁷ For atti-

tudes and behaviors that are positively valued in a society, on the other hand—for example, voting, owning a library card, holding liberal racial views—less reporting may be considered better reporting (e.g., Cahalan, 1968; Clausen, 1968; Freeman, 1953; Parry and Crossley, 1950; Weiss, 1968). Accordingly, we can ask two questions of the data from the present study: Do informed consent procedures have any significant or consistent effects on response; and if so, do they make for better reporting, or worse?

As in the analysis of response rate to the

¹⁷ Blair et al. (1977) cite studies by Cannell and Fowler (1963); Clark and Wallin (1964); Clark and Tift (1966); David (1962); DeLamater and MacCor-

quodale (1975); Ellis (1947); Johnson and DeLamater (1976); Kahn (1952); Kinsey et al. (1948); Knudsen et al. (1967); Levinger (1966); Locander et al. (1976); Mudd et al. (1961); Poti et al. (1962); Sarason (1956; 1957; 1959); Thorncike et al. (1952); U. S. National Center for Health Statistics (1971); Wallin and Clark (1958); Yaukey et al. (1965); and Young (1969).

Table 9. Nonresponse ^a to Sensitive ^b Questions, by Signature Condition

Question	Percent Nonresponse				Signifi- cance ^c
	Signed Before	Signed After	Not Asked	Refused	
1. Ever smoked marijuana 3 time a week or more? ^d	8.7% (104)	7.1% (98)	5.9% (101)	0.0% (7)	n.s.
2. Number of pipes, joints per time? ^e	5.8 (104)	5.1 (98)	5.9 (101)	0.0 (7)	n.s.
3. Engaged in petting or kissing in past month?	4.8 (392)	6.9 (406)	4.9 (449)	18.9 (74)	<.05
4. How often? ^d	7.5 (266)	6.5 (276)	7.7 (323)	24.4 (45)	<.05
5. Petted, kissed in past 24 hours? ^d	2.3 (266)	3.3 (276)	5.3 (323)	4.4 (45)	n.s.
6. Intercourse in past month?	6.1 (392)	8.9 (406)	8.2 (449)	27.0 (74)	<.05
7. How often? ^d	9.0 (244)	9.7 (238)	10.1 (287)	19.4 (31)	n.s.
8. Intercourse in past 24 hours? ^d	3.7 (244)	5.0 (238)	4.5 (287)	12.9 (31)	n.s.
9. Masturbated in past month?	8.2 (392)	11.6 (406)	10.2 (449)	28.4 (74)	<.05
10. How often? ^d	0.0 (24)	12.9 (31)	8.6 (35)	0.0 (3)	n.s.
11. Masturbated in past 24 hours? ^d	0.0 (24)	9.7 (31)	5.7 (35)	0.0 (3)	n.s.
12. Earned income in past year?	9.4 (392)	9.4 (406)	10.0 (449)	21.6 (74)	<.05

^a "Nonresponse" includes refusal, don't know, and not asked.

^b All questions on which more than 3% of responses were coded as missing data are included in this table. With one exception, they come from the "sensitive, survey-specific, factual" question category. The exception is the question on income, which is general rather than survey-specific.

^c Significance levels are based on chi-square.

^d For these questions we give the conditional probability of refusing to answer, among those who answered the filter question affirmatively.

^e On this particular question, 16 of the 17 missing responses are attributable to interviewer omissions rather than respondent refusals.

questionnaire as a whole, the method of analysis used is Multiple Classification Analysis, which is particularly well suited to the categorical form of the independent variables. The assumption of an additive model, on which multiple classification analysis is based, is supported by the fact that fewer than one in twenty of the two-way interactions among independent variables reach statistical significance. In examining the effect of each independent variable, we control not only for the other two but also for age, sex, and education, which influence many of the attitudes and behaviors in question.

For the analysis of response quality, we selected samples of nonsensitive attitude

and behavioral questions, but have included virtually every question, whether about attitudes or behavior, that could be construed as sensitive. The effects of all three independent variables are summarized in Table 13; more detailed tables are available from the author on request.

The evidence for the effect of information about content on response quality is mixed. Only five comparisons (out of 43) between experimental conditions are statistically significant. On two sensitive, survey-specific attitude items, the vague, brief introduction results in better response quality (more reported negative affect, less reported happiness). On three behavioral items (income, masturbation,

Table 10. Variations in Nonresponse,^a by Experimental Condition

Experimental Condition	Percent with No Nonresponse to Any of 4 Items	Mean Number of Nonresponses, among Those with Any
Content:		
Long	81% (659)	2.08 (N=127)
Short	82 (662)	1.85 (N=118)
Significance of difference ^b	$p=.47$	$p>.10$
Confidentiality:		
No mention	77 (457)	2.04 (N=104)
Qualified	81 (430)	1.94 (N= 82)
Absolute	87 (434)	1.88 (N= 59)
Significance of difference ^b	$p<.01$	$p(F)=.67$
Signature:		
Before	85 (392)	1.93 (N= 58)
After	81 (406)	1.94 (N= 77)
Not asked	82 (449)	1.83 (N= 82)
Refused	62 (74)	2.54 (N= 28)
Significance of difference ^{b, c}	$p<.01$	$p(F)=.05$

^a Based on an index of four items (see text).

^b Based on chi-square for proportions and t-tests or one-way analysis of variance for means.

^c When those who refused to sign are excluded from the analysis, the difference among conditions is not significant.

probability of having smoked marijuana three times a week or more), the long introduction results in higher estimates (i.e., better quality) than the short.¹⁸ Even if one chooses to ignore statistical significance and to look instead at consistency of response, no clear pattern is discernible.

¹⁸ The problem with using consistency of response instead of statistical significance as a guide to response effects is selection of the items to be examined for consistency. If, for example, one decided to use only the most sensitive items on the questionnaire, as defined by item nonresponse rate, one would look at items VIII. 1, 2, 3, 4, 5, 6, 7, 8, 11, 12, and V. 2. (Because so few respondents answered the question, we did not include probability of masturbating during the last 24 hours in the analysis shown in Table 13.) If the analysis were restricted to these items, conclusions concerning the effects of content and signature would not change, but since six of ten comparisons (instead of an expected 3.3) favor the absolute confidentiality condition, one might be tempted to modify one's conclusions concerning the effect of confidentiality on response quality.

If the pool of items were restricted even further, to

Table 11. Nonresponse^a to Sensitive^b Questions, among Those Refusing to Sign Before and Those Refusing to Sign After the Interview

Question	Percent Nonresponse	
	Refused to Sign Before	Refused to Sign After
1. Ever smoked marijuana 3 times a week or more?	0.0% (5)	0.0% (2)
2. Number of pipes, joints per time?	0.0 (5)	0.0 (2)
3. Engaged in petting or kissing in last month?	13.6 (44)	26.7 (30)
4. How often?	23.3 (30)	26.7 (15)
5. Petted, kissed in last 24 hours?	6.7 (30)	0.0 (15)
6. Intercourse in past month?	27.3 (44)	26.7 (30)
7. How often?	12.5 (16)	26.7 (15)
8. Intercourse in past 24 hours?	6.3 (16)	20.0 (15)
9. Masturbated in past month?	27.3 (44)	30.0 (30)
10. How often?	0.0 (2)	0.0 (1)
11. Masturbated in last 24 hours?	0.0 (2)	0.0 (1)
12. Earned income in past year?	13.6 (44)	33.3 (30)

^a "Nonresponse" includes refusal, don't know, and not asked.

^b All questions on which more than 3% of responses were coded as missing data are included in this table.

the four most sensitive questions asked of the entire sample (items V.2 and VIII. 1, 4, and 7 in Table 13), three of four comparisons would favor the detailed (long) introduction and the absolute confidentiality condition, but no clear trend would emerge for the signature variable.

Which of these three sets of questions is most appropriate for drawing conclusions is by no means clear, unfortunately. Restricting interpretations to those differences which are statistically significant is not necessarily the answer, either, because, at least in the case of confidentiality, we have reason to believe that the effects we observed are minimum effects.

Table 12. Refusal to Sign Consent Form, by Selected Demographic Characteristics

Demographic Characteristic	Percent of All Those Asked to Sign Who Refused	
Race:		
White	8.7%	(807)
Black	5.6	(54)
Sex:		
Male	8.6	(349)
Female	8.4	(523)
Age:		
18-25	2.6	(156)
26-35	7.7	(207)
36-45	9.8	(123)
46-55	7.1	(127)
56-65	11.3	(133)
66+	14.3	(126)
Education:		
11 years or less	11.6	(275)
12 years	6.9	(331)
13-15 years	5.7	(140)
16 years or more	8.9	(124)

Variations in confidentiality produce only three statistically significant differences in response quality: one in the category of nonsensitive, general attitude questions; one in the category of sensitive, survey-specific attitude questions; and one in the category of nonsensitive, general factual questions. In two of these comparisons, the condition in which no mention is made of confidentiality produces the best data (lower scores on the Bradburn Positive Affect Scale, less likelihood of giving one's occupation as professional or managerial); in the third, it is the qualified confidentiality condition which produces the best data (less reported satisfaction with the neighborhood).

These three instances do not speak compellingly for or against any confidentiality condition, nor is any consistent pattern of response apparent if one considers the entire pool of items in Table 13. However, if one looks only at the most sensitive items, as defined by item nonresponse (see fn. 18), one finds that on six of them, respondents in the absolute confidentiality condition produce the highest estimates, with a tie between those in the absolute and the no-mention conditions on an

additional item. Similarly, the condition in which the respondent is promised absolute confidentiality produces the highest estimates on three of the four most sensitive questions asked of the entire sample. Thus, there is at least the suggestion that a promise of absolute confidentiality enhances the quality of response to sensitive questions, over and above its effect on item nonresponse.

As in the case of content and confidentiality, very few of the comparisons involving the signature variable reach statistical significance. The five that do, however, all involve sensitive, survey-specific behavioral questions, and all of them indicate that respondents who sign the consent form, either before or after the interview, report higher frequencies of such behavior than respondents who are not asked to sign.

One reason for this result is that respondents who refused to sign the consent form, and who therefore are properly excluded from the analysis reported in the paragraph immediately above, tend to give the lowest estimates of sensitive behavior.¹⁹ In other words, people who sign the consent form produce better data than those not asked to sign because those who report the lowest frequencies of sensitive behavior also refuse to sign the consent form and are, therefore, excluded from the analysis.

However, the estimates of threatening behavior given by those who sign the consent form after the interview tend to be higher than those given by respondents who sign before. Since the latter group shows little difference from the group not asked to sign at all, even though the request for a signature screens out some underestimators, we conclude that asking for a signature before the interview has a sensitization effect. Respondents who

¹⁹ Analyses of the effects (on item response rate and response quality) of varying information about content and assurance of confidentiality are based on all those interviewed—that is, they include the 74 respondents who refused to sign the consent form. (The results for content and confidentiality do not change in any significant way if these 74 respondents are excluded.) People who refused to sign the consent form are, of course, excluded from comparisons among the signature conditions. Their item nonresponse rate is shown separately in Table 9.

Table 13. Responses to Different Types of Questions, as Function of Content, Confidentiality, and Request for Signature *

Type of Question	Content	Confidentiality	Signature
I. Effect on Sensitive, General, Attitude Questions			
1. Social desirability scale	S ^b	N ^a	B ^d
2. Satisfaction with financial situation	S	N	B
II. Effect on Nonsensitive, General, Attitude Questions			
1. Satisfaction with house	L	Q	=
2. Satisfaction with neighborhood	S	Q*	B
III. Effect on Sensitive, Survey-Specific, Attitude Questions			
1. Positive Affect Scale	=	N*	B
2. Negative Affect Scale	S*	C	A
3. 22-Item Scale	S	Q	N
4. Reported happiness	S*	Q	A
IV. Effect on Nonsensitive, Survey-Specific, Attitude Questions			
1. Importance of leisure	L	N	B
2. Satisfaction with leisure	S	N	A
V. Effect on Sensitive, General, Factual Questions			
1. Number of sources of income	L	C	B
2. Amount of earned income	L*	Q	A
VI. Effect on Nonsensitive, General, Factual Questions			
1. Number of years of school	S	C	N
2. Probability of giving occupation as professional or managerial	S	N*	=
VII. Effect on Nonsensitive, Survey-Specific, Factual Questions			
1. Probability of going to a restaurant	=	=	N
2. Probability of bowling	L	=	=
3. Probability of swimming	S	C	A
4. Frequency of giving a party	L	Q	A
5. Frequency of being with relatives	L	C	B
VIII. Effect on Sensitive, Survey-Specific, Factual Questions			
1. Probability of petting, kissing in past month	L	C	N
2. Petting frequency	S	C	A
3. Petting in past 24 hours	S	Q	A
4. Probability of intercourse in past month	=	C	N
5. Frequency of intercourse	S	C	A
6. Intercourse in past 24 hours	S	=	A*
7. Probability of masturbation in past month	L*	C	=
8. Frequency of masturbation	L	C	A
9. Probability of smoking marijuana	S	N	=

Table 13. Continued

Type of Question	Content	Confidentiality	Signature
VIII. Effect on Sensitive, Survey-Specific, Factual Questions—(Continued)			
10. Conditional probability of smoking in last year	L	C	B
11. Conditional probability of smoking 3 times per week or more	L*	N	B*
12. Number of pipes, joints smoked	L	N	B
13. Number of close friends who smoke	S	N	A*
14. Number of times drunk last year	L	N	A
15. Number of friends drunk last year	S	N	B*
16. Probability of drinking liquor	S	N	B
17. Conditional probability of drinking last year	S	Q	A
18. Frequency of drinking	S	C	A
19. Number of drinks per time	L	C	N
20. Probability of drinking beer	S	=	N
21. Conditional probability of drinking beer	S	Q	A
22. Frequency of drinking beer	L	N	A
23. Number of beers per time	L	Q	A*
24. Gambling Scale score	S	N	A

^a Respondents who refused to sign the consent form are excluded from the analysis.

^b "L" means that respondents given the long introduction produce better responses; "S" means that those given the short introduction do; an "equal" sign means there are no differences. For a definition of "better," see text.

^c "C" means those given an absolute assurance of confidentiality give the best responses; "N" means those to whom confidentiality is not mentioned do; "Q," that respondents in the qualified confidentiality condition do. An "equal" sign signifies a tie between the absolute and another condition.

^d "B" means those asked to sign before give the best responses; "A," that those asked afterwards do; "N" means that those not asked for a signature do. An "equal" sign signifies a tie between two conditions—"not asked" and "after" on VII, 2 and VIII, 7; "before" and "after" on the remaining items.

* $p(F) < .05$.

sign the consent form before the interview are more likely to underestimate socially undesirable behavior than if they had not been asked to sign at all, or if they had been asked to sign afterwards.

So far, we have examined response tendencies produced by the three independent variables considered one at a time, and have concluded that these are, for the most part, small. But ethical considerations and statutory requirements call for incorporating not one, but at least two and perhaps all three variables into survey introductions in order to secure the informed consent of the respondent. What, then, is the cumulative effect on response quality of all three independent variables examined in this study?

One approach to answering this question is to contrast the two extreme conditions already described. One is the standard survey introduction, in which the re-

spondent is ordinarily given only vague, general information about the content of the interview, is assured that his replies will be held in strict confidence, and is not asked to sign a consent form. The other is the condition that, it might be argued, most nearly assures the informed consent of survey respondents. It provides detailed information about the content of the interview, qualifies the assurance of confidentiality, and asks the respondent to document his understanding and agreement by signing a consent form before being interviewed. Under these contrasting assumptions, what effects on response tendencies can be discerned?

In order to answer this question, we examined responses in the two experimental conditions (out of 18) which represent the two sets of contrasting assumptions above. Because the numbers of cases are now very small (81 in the stan-

dard and 66 in the informed consent condition), we did not attempt to control for other variables.

Some five differences between the two conditions are statistically significant. Of these, three indicate better responses on the part of those in the informed consent condition, though only one of the three is a response to a sensitive item.

If we look at the pattern of responses rather than at statistical significance, the standard introduction appears to yield better data on sensitive, survey-specific attitude questions, whereas the introduction involving informed consent yields better data on 14 of 19 sensitive behavioral items on which any difference is discernible. Because of the absence of controls for age, sex, or education, we do not wish to overemphasize these findings; in fact, a different approach²⁰ to this question suggests that the standard introduction produces data of better quality on questions of both types.

Summary and Conclusions

This study was designed to provide information to guide policy. What are the implications of its findings?

Four of those findings deserve emphasis. First, of the three experimental variables—content, confidentiality, and signature—only the request for a signature affected the response rate to the questionnaire as a whole. About 8% of those asked to sign a consent form refused to do so, although they were willing to be interviewed.

Second, those who refused to sign a consent form were also much more likely to refuse to answer individual questions within the interview and to provide poorer data when they did answer. Thus, paradoxically, the estimates obtained from a survey on sensitive topics may be improved if those who refuse to sign are ex-

cluded from the sample. However, this is true only if the request for a signature comes after the interview. For those asked to sign before, the request appeared to have a sensitization effect, so that these respondents were more likely to underestimate socially undesirable behavior than if they had not been asked to sign at all or if they had been asked to sign afterwards.

Since the request for a signature appears to function largely as another sensitive question, the requirement that researchers obtain a signature to document consent seems unnecessarily burdensome. The same protection is afforded respondents by the right to refuse the interview, or to refuse to answer particular questions within the interview.²¹ However, if a signature is required, it should be obtained after the interview rather than before, in order not to jeopardize further the quality of response.

The third major finding is that, even though it was impossible to discern any effect of confidentiality on overall response rate to the interview, assuring respondents of absolute confidentiality had a small but consistent effect on willingness to answer individual questions. Nonresponse rates for sensitive questions were consistently and sometimes significantly lower when people were told that their replies would be held in confidence. Furthermore, there is at least the suggestion that a promise of confidentiality enhances the quality of response to the most sensitive items on the interview. Thus, if respondents can be assured of the confidentiality of their replies, responses to sensitive questions will benefit. But such assurances must be meaningful; they cannot be given lightly.

Finally, since a more detailed, informative, and truthful introduction adversely affects neither overall response rate nor responses to individual questions, there appears to be no reason to withhold such information from respondents.

²⁰ Since an additive model adequately describes the effects of variations in information, assurance of confidentiality, and signature on response quality, it is possible to sum the adjusted deviations associated with separate effects in order to obtain estimated values for the two contrasted conditions. It is not possible, however, to compute the significance of the difference between these estimates.

²¹ This conclusion of course pertains only to those situations in which a signature is obtained from the same person who is also to be interviewed, and not to situations in which a signature may be required in order to document a parent's consent to interview a child, or to give the researcher access to the respondent's personal records.

All of these findings, it should be emphasized, derive from one type of survey only. It is possible that certain kinds of questions, asked of certain specialized categories of respondents, might interact with the independent variables to produce results other than those reported here. For example, if welfare clients were asked about their income, refusals under several of the experimental conditions might be higher than those in the present study; the same is true if employees of a large corporation were asked about their drinking habits. One study cannot hope to answer all such questions; the present one specifies what is likely to happen in a general population survey dealing with generally sensitive content.

Just what the larger implications of the findings reported here are, is far from clear. One may conclude optimistically that none of the elements of informed consent, except the request for a signature, has sizable effects on the response rate to surveys or the quality of response, and that therefore ethical imperatives do not conflict with practical considerations. Or, one may take the more pessimistic view that respondents simply do not attend to what they are told, deciding whether to participate or not on grounds entirely extraneous to those which are experimentally varied here. Two-thirds of those who refused the interview, for example, had heard no part of the introduction whatsoever. On this more pessimistic view, the findings reported indicate only that even the procedures used in this survey fail to assure truly informed consent on the part of participants in social research.

REFERENCES

- Aronson, E., and J. M. Carlsmith
1969 "Experimentation in social psychology." Pp. 1-79 in G. Lindzey and E. Aronson (eds.), *Handbook of Social Psychology*. 2nd ed. Vol. 2. Reading, Ma.: Addison-Wesley.
- Ash, P., and E. Abramson
1952 "The effect of anonymity on attitude questionnaire response." *Journal of Abnormal and Social Psychology* 47:722-3.
- Barber, Bernard
1973 "Experimentation with human beings: another problem of civil rights?" *Minerva* 2: 415-9.
- Barber, Bernard, J. J. Lally, J. L. Makarushka, and D. Sullivan
1973 *Research on Human Subjects*. New York: Russell Sage.
- Becker, G., and D. A. Bakal
1970 "Subject anonymity and motivational distortion in self-report data." *Journal of Clinical Psychology* 26:207-9.
- Blair, E., S. Sudman, N. M. Bradburn, and C. Stocking
1977 "How to ask questions about drinking and sex." *Journal of Marketing Research* 14: 316-21.
- Blumberg, Herbert H., Carolyn Fuller, and A. Paul Hare
1974 "Response rates in postal surveys." *Public Opinion Quarterly* 38:113-23.
- Bradburn, N. M., S. Sudman, E. Blair, and C. Stocking
1978 "Question threat and response bias." *Public Opinion Quarterly*. In press.
- Cahalan, D.
1968 "Correlates of respondent accuracy in the Denver validity survey." *Public Opinion Quarterly* 32: 607-21.
- Cannell, C. F., and F. J. Fowler
1963 "A comparison of a self-enumerative procedure and a personal interview: a validity study." *Public Opinion Quarterly* 27:250-64.
- Clark, A. L., and P. Wallin
1964 "Accuracy of husbands' and wives' reports of frequency in marital coitus." *Population Studies* 18:165-73.
- Clark, J. P., and L. L. Tift
1966 "Polygraph and interview validation of self-reported deviant behavior." *American Sociological Review* 31:516-23.
- Committee on Federal Agency Evaluation Research
1975 *Protecting Individual Privacy in Evaluation Research*. Assembly of Behavioral and Social Sciences, National Research Council, National Academy of Sciences, Washington, D. C.
- Clausen, A. R.
1968 "Response validity: vote report." *Public Opinion Quarterly* 32:588-606.
- David, M.
1962 "The validity of income reported by a sample of families who received welfare assistance during 1959." *Journal of the American Statistical Association* 57:680-5.
- DeLamater, J., and P. MacCorquodale
1975 "The effects of interview schedule variations on reported sexual behavior." *Sociological Methods and Research* 4:215-36.
- Ellis, A.
1947 "Questionnaire versus interview methods in the study of human love relationships." *American Sociological Review* 12:541-53.
- Erdos, P., and J. Regier
1977 "Visible and disguised keying on questionnaires." *Journal of Advertising Research* 17:13-8.

- Fischer, R. P.
1946 "Signed versus unsigned personal questionnaires." *American Psychologist* 5:315.
- Freeman, H. E.
1953 "A note on the prediction of who votes." *Public Opinion Quarterly* 17:288-92.
- Fuller, C.
1974 "Effect of anonymity on return rate and response bias in a mail survey." *Journal of Applied Psychology* 59:292-6.
- Gray, Bradford H.
1975 *Human Subjects in Medical Experimentation*. New York: Wiley.
- Haberman, Paul, E. Josephson, A. Zanes, and J. Elinson
1972 "High school drug behavior: a methodological report on pilot studies." Pp. 103-21 in S. Einstein and S. Allen (eds.), *Proceedings of the First International Conference on Student Drug Surveys*. Farmingdale, N. Y.: Baywood.
- Hamel, L., and H. G. Reif
1952 "Should attitude questionnaires be signed?" *Personnel Psychology* 5:87-92.
- Hauck, M., and Michael Cox
1974 "Locating a sample by random digit dialing." *Public Opinion Quarterly* 38:253-6.
- Johnson, Weldon T., and John D. DeLamater
1976 "Response effects in sex surveys." *Public Opinion Quarterly* 40:165-81.
- Jourard, S. M.
1964 *The Transparent Self*. Princeton: Van Nostrand.
1968 *Disclosing Man to Himself*. Princeton: Van Nostrand.
- Jourard, S. M., and R. Friedman
1970 "Experimenter-subject 'distance' and self-disclosure." *Journal of Personality and Social Psychology* 15:278-82.
- Kahn, R. L.
1952 *A Comparison of Two Methods of Collecting Data for Social Research: The Fixed-Alternative Questionnaire and the Open-Ended Interview*. Ph.D. dissertation, Department of Sociology, University of Michigan.
- King, F. W.
1970 "Anonymous versus identifiable questionnaires in drug usage surveys." *American Psychologist* 25:982-5.
- Kinsey, A. C., W. B. Pomeroy, and C. E. Martin
1948 *Sexual Behavior in the Human Male*. Philadelphia: Saunders.
- Knudsen, D. D., H. Pope, and D. P. Irish
1967 "Response differences to questions on sexual standards: an interview-questionnaire comparison." *Public Opinion Quarterly* 31:290-7.
- Kosen, D., C. Kitchen, M. Kochen, and D. Stodolsky
1970 "Psychological testing by computer: effect on response bias." *Educational and Psychological Measurement* 30:803-10.
- Levinger, G.
1966 "Systematic distortion in spouses' reports of preferred and actual sexual behavior." *Sociometry* 29:291-9.
- Locander, W. B., S. Sudman, and N. M. Bradburn
1976 "An investigation of interview method, threat and response distortion." *Journal of the American Statistical Association* 71: 269-75.
- Mason, W. S., R. J. Dressel, and R. K. Bain
1961 "An experimental study of factors affecting response to a mail survey of beginning teachers." *Public Opinion Quarterly* 25: 296-9.
- Mitchell, W., Jr.
1939 "Factors affecting the rates of return on mailed questionnaires." *American Statistical Association Journal* 34:683-92.
- Mudd, E. H., M. Stein, and H. E. Mitchell
1961 "Paired reports of sexual behavior of husbands and wives in conflicted marriages." *Comprehensive Psychiatry* 2:149-56.
- National Data Program for the Social Sciences
1976 *Codebook for the Spring 1976 General Social Survey*. New York: National Opinion Research Center.
- Nejelski, Paul (ed.)
1976 *Social Research in Conflict with Law and Ethics*. Cambridge, Ma.: Ballinger.
- Parry, H. J., and H. M. Crossley
1950 "Validity of responses to survey questions." *Public Opinion Quarterly* 14: 61-80.
- Poti, S. J., B. Chakraborti, and C. R. Malaker
1962 "Reliability of data relating to contraceptive practices." Pp. 51-65 in Clyde V. Kiser (ed.), *Research in Family Planning*. Princeton: Princeton University Press.
- Privacy Protection Study Commission
1977 *Personal Privacy in an Information Society*. Washington, D.C.: U.S. Government Printing Office.
- Rosen, N. A.
1960 "Anonymity and attitude measurement." *Public Opinion Quarterly* 24:675-9.
- Sarason, I. G.
1956 "Effect of anxiety, motivational instructions, and failure on serial learning." *Journal of Experimental Psychology* 51:253-60.
1957 "Effect of anxiety and two kinds of motivating instructions on verbal learning." *Journal of Abnormal and Social Psychology* 54:166-71.
1959 "Relationships of measures of anxiety and experimental instructions to word association test performance." *Journal of Abnormal and Social Psychology* 59:37-42.
- Sudman, S., N. M. Bradburn
1974 *Response Effects in Surveys: A Review and Synthesis*. Chicago: Aldine.
- Sudman, S., N. M. Bradburn, E. Blair, and C. Stocking
1977 "Modest expectations: the effect of interviewer expectations on responses to threatening questions." Xerox. New York: National Opinion Research Center.
- Thorndike, R. L., E. Hagen, and R. A. Kemper
1952 "Normative data obtained in the house-to-house administration of a psychosomatic inventory." *Journal of Consulting Psychology* 16:257-60.

- U.S. National Center for Health Statistics
 1971 Effect of Some Experimental Interviewing Techniques on Reporting in the Health Interview Survey. Vital and Health Statistics, Series 2, No. 41. Washington, D.C.: U.S. Government Printing Office.
- Wallin, P., and A. Clark
 1958 "Cultural norms and husbands' and wives' reports of their marital partners' preferred frequency of coitus relative to their own." *Sociometry* 21:247-54.
- Warwick, Donald P.
 1975 "Social scientists ought to stop lying." *Psychology Today* 8 (February):38 ff.
- Weiss, C. H.
 1968 "Validity of welfare mothers' interview responses." *Public Opinion Quarterly* 32: 622-32.
- Wildman, Richard C.
 1977 "Effects of anonymity and social setting on survey responses." *Public Opinion Quarterly* 41:74-9.
- Yaukey, D., B. J. Roberts, and W. Griffiths
 1965 "Husbands' vs. wives' responses to a fertility survey." *Population Studies* 19:29-43.
- Young, B. A.
 1969 The Effects of Sex, Assigned Therapist or Peer Role, Topic Intimacy, and Expectations of Partner Compatibility on Dyadic Communication Patterns. Ph.D. dissertation, University of Southern California.

DETERMINANTS OF JUVENILE COURT DISPOSITIONS: ASCRIPTIVE AND ACHIEVED FACTORS IN TWO METROPOLITAN COURTS*

LAWRENCE E. COHEN

University of Illinois, Urbana

JAMES R. KLUEGEL

University of California, Riverside

American Sociological Review 1978, Vol. 43 (April):162-176

This paper examines the impact of stereotypical and discriminatory factors on the severity of dispositions accorded juveniles in two courts, Denver and Memphis, which differed in their orientation to juvenile justice and in regional location. Because the conclusions of earlier research on juvenile justice decision making are affected by inadequate data-analytic techniques, this study analyzes multivariate relationships among qualitative variables using Goodman's method of log linear analysis to investigate possible sources of bias in the severity of disposition for 6,894 male juveniles. Little support is found for the argument that race or social class bias directly affects the dispositions in these two courts. The implications of this research for understanding the prior pattern of contradictory findings and for the general issue of bias in the juvenile justice system are discussed.

According to a number of recent assessments, the contemporary American juvenile justice system incorporates biases which virtually ensure that specific youths, particularly minority group members and those of lower socioeconomic status, will be the objects of discriminatory treatment (Cicourel, 1968; Platt, 1969;

Martin, 1970; Schur, 1973). This charge takes several forms, but in essence the contention is that certain youths are more likely to be accorded harsh treatment, not necessarily due to the nature of the offenses for which they are charged, or due to a prior history of delinquency involvement, but rather because they fit the preconceived notions or stereotypes of the delinquent which court officials have formulated. Such youths are thought to be systematically treated more harshly by social control agents than are other juveniles (presumably, those who do not fit the control agents' stereotype of a delinquent), who may often be released without official court intervention.

Martin (1970:3-4) for example, contends, that:

* This project was supported by Grant No. 72-SS-99-006, awarded to the Criminal Justice Research Center (Albany, New York) by the Statistics Division, National Criminal Justice Information and Statistics Service, Law Enforcement Assistance Administration, U.S. Department of Justice. Views or opinions stated in this paper are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice. The authors wish to thank Eliot Smith for helpful comments and suggestions on an earlier draft of this paper.

In effect, the juvenile justice labeling process works to single out adolescents from groups culturally alien to those in power. Those singled out, because of their powerlessness, are ill-equipped to stop the process or to intervene in it effectively to prevent themselves from having various and sundry tags imposed upon them by police, judges, probation officers, psychiatrists, and others who are employed as agents of the juvenile justice system.

Martin (1970:4) maintains that once the child's case is brought to the attention of the juvenile court, it is handled by probation officers and judges whose social and cultural characteristics are decidedly middle class, and that this fact makes it difficult for these functionaries to be objective when judging the behavior of lower-class youths.

Edwin Schur (1973:121) expounds further on the process by which stereotyping is likely to work against youths from certain disadvantaged segments of our society, regardless of the offenses with which they are charged:

The philosophy of the juvenile court—with its thoroughgoing social investigation of the alleged delinquent, and its relative lack of concern with the particular offense—virtually ensures that stereotypes will influence judicial dispositions.

Schur (1973:125–6) subsequently identifies those juveniles whom he believes to be the most likely recipients of differential treatment from agents of the juvenile court because they are believed to be “delinquency-prone:”

In our society, lower-class children more than middle-class ones, black children more than white ones, and boys more than girls, face high probabilities (i.e., run a special “categorical risk” in the actuarial sense) not only of engaging in rule-violation in the first place, but also of becoming enmeshed in official negative labeling processes.

While few would deny that lower-class and minority group members are overrepresented at the various stages of the criminal justice process, the issue is whether that overrepresentation reflects stereotyping or bias in the system, or whether it is attributable to such group members' greater proportional involvement in offenses or possession of prior criminal or

delinquent records which might justify severe treatment. This issue is further complicated in the area of juvenile court decision making by differences in orientation followed by various juvenile courts around the country, and by the methodological inadequacies of the studies which attempt to assess the criteria upon which dispositions are accorded in these courts.

The findings of the more rigorous empirical studies have been less than conclusive with respect to the role of stereotyping and discrimination in juvenile court decision making.¹ For example, Terry's (1967) study of the Racine, Wisconsin Juvenile Court indicates that no class or race effects are evident when statistical controls are applied to his data. The severity of disposition in the Racine Court was most strongly related to the juvenile's prior offense record and the seriousness of the offense for which the youth was charged. Arnold's (1971) study of a middle-sized southern city, on the other hand, suggests that evidence of discriminatory treatment against blacks was apparent both before and after controls were added to his analysis, but that no substantial differences in treatment could be observed among the various social classes. Finally, Thornberry's (1973) analysis of dispositions in the Philadelphia Juvenile Courts reports that race and class significantly affect the treatment meted out, with black and lower-class male youths being systematically accorded more severe treatment, even with controls for prior record and seriousness of offense added to the analysis.

One possible explanation for this pattern of contradictory findings may lie in the approaches to juvenile justice taken by different courts. Different courts might well adhere to different orientations or philosophies concerning the treatment of juvenile offenders. For example, due to the impact of recent Supreme Court decisions, some juvenile courts have developed systems which pay greater attention

¹ Discrimination or bias is generally considered present in these studies if minorities and lower-class juveniles are more likely to have had their offenses handled officially by the court and/or more likely to have been incarcerated in a training institution than were nonminorities and nonlower-class youths.

to due process guarantees for juveniles (see, for example, in re Gault, 387 U.S. 1, 1967; *Kent v. United States*, 383 U.S. 541, 1966; and in re Winship, 397 U.S. 358, 1970). It has been suggested that bias is less apt to occur in courts committed to the due process model, for, as Herbert Packer (1966: 39) has noted, this model of processing has the effect of:

"Judicializing" each stage of the criminal process, of enhancing the capacity of the accused to challenge the operation of the process, and of equalizing the capacity of all persons to avail themselves of the opportunity for challenge so created.

Hirschi (1975: 191), on the other hand, has recently characterized the traditional therapeutic juvenile court model and summarized the views of many regarding this model's hypothesized effect on juvenile court processing:

As everyone knows, the (traditional) juvenile justice system was explicitly constructed to give the kindly agents of the state a relatively free hand in dealing with the problems of children. This system was authorized to take into account the needs of the child, his or her probable future behavior, and so on through a long list of considerations that would seem to allow or even require bias or discrimination on the part of officials.

While the possible effect of court differences in the approach to juvenile justice should be considered in accounting for the contradictory findings of previous studies, we must consider as well the possibility that differences in the techniques of analysis have contributed to the variation in the reported results. In our judgment, none of these studies has utilized a data-analytic technique which allows for an adequate treatment of the issue of bias in the juvenile court.

As we have previously noted, the issue of bias as it is usually treated in the empirical literature involves the question of whether extra-legal factors (i.e., race, class, etc.) lead to differential treatment of juvenile cases when relevant legal factors are controlled for (i.e., seriousness of offense, prior record, etc.). To adequately treat this issue, relevant legal variables must be introduced simultaneously into the analysis. An analysis should also allow

estimation of the magnitude of the bias, should it be demonstrated that bias is present. In both of these respects the literature cited above is in our judgment deficient, or at least ambiguous. Terry's (1967) multivariate findings, for example, were obtained through the applications of control variables introduced one or two at a time, but not all at once. The simultaneous introduction of several statistical controls may substantially alter the results of one's analysis. In addition, Terry used a single "average" Kendall's tau, correlation coefficient to summarize the magnitude of the relationships occurring within the subclassifications of his control variables. Thus the average tau he obtained may have been insensitive to possible interaction effects within the various levels of the control variable utilized. Arnold's (1971) multivariate findings, on the other hand, are largely dependent on what he terms a "total consideration score" compiled from his independent variables. However, the procedure by which he constructs this multivariate index is of questionable validity.² Finally, Thornberry's (1973) failure to assess the magnitude of the presumed relationships in his data has apparently led him to conclusions which are not justified. Wellford (1975:338-9) effectively demonstrates that when a measure of association (instead of the percentage differences employed by Thornberry) is utilized to assess the strength of the relationships reported by Thornberry, "the data reflect the minimal contribution of race and SES to criminal justice decision making." Thus, methodological inadequacies and failures to consider the effect of court differences in previous studies call for further empirical assessment of the nature and extent of bias in the treatment of juvenile offenders.

This paper is an attempt to address the issue of stereotyping and discrimination in our juvenile justice system by:

² Arnold's "Total Consideration Score" is an ad hoc procedure employed to control simultaneously for the effects of several independent variables on juvenile dispositions. Arnold presents no logical justification for the weighting procedure used in compiling this score. In addition, the cell frequencies in the devised categories are so small that they lead to questions regarding the stability of these findings under replication.

- (1) utilizing a more appropriate and rigorous system for the analysis of multivariate relationships among qualitative variables, including statistical interactions (Goodman's method of log linear analysis; see Goodman, 1972 and 1973, for a general discussion of log linear analysis, and Burke and Turk, 1975, for an application of this method to the sentencing of adult offenders); and
- (2) comparing treatment accorded in two courts differing in orientation to juvenile justice, which allows us to assess intercourt differences.

The Data

The data employed in this study were gathered from completed case history records compiled for all male juveniles (excluding dependency, neglect, and traffic cases) referred to the Denver (Colorado), and Memphis (Tennessee) juvenile courts between January 1, and December 31, 1972—a total of 2,465 cases from Denver and 4,429 cases from Memphis.³ The comparability of the data permits the assessment of court variations in the handling of juvenile offenders. Both courts have their own judges who deal exclusively with juvenile cases and, in addition, the courts have comparable administrative structures to handle the variety of functions which they perform (for a complete description of the administrative structures of these courts and a preliminary analysis of these and other data see Cohen, 1975).

Though there are similarities in the structure of these two respective juvenile courts, there is an important difference in the overall orientation of the court per-

sonnel within these two centralized bureaucracies toward the youths who come to their attention. The Denver facility is considerably more due process oriented.⁴ The personnel of the Denver facility exhibited an overriding concern for and attention to procedural rights of youths, and the court functions were structured in ways designed to safeguard these rights. The overall orientation of the Memphis facility, on the other hand, resembles more closely the traditional therapeutic model for juvenile courts (i.e., one in which the proper course of action should be taken "in the best interest of the child"). Therapeutic concerns were voiced more often and more emphatically by personnel of the Memphis Court than by personnel of the Denver Court.⁵ As we

⁴ The Denver Juvenile Court is certainly one of the most due process oriented juvenile tribunals in the country; it adheres largely to the model practiced in adult criminal courts. For example, the Denver court requires that the police adhere to a probable cause standard for arrest before any child can be adjudicated as a delinquent. If a child is to be detained for any length of time exceeding 24 hours prior to his adjudication, the Denver Juvenile Court permits the posting of bail. Whenever a child's case is to be adjudicated before a formal court hearing, the court requires that the youth be represented by an attorney and permits (and even encourages) the use of plea bargaining during the adjudication process. Finally, if the attempt at plea bargaining is unsuccessful, the court permits the child (or his parents, guardian, or attorney) to choose between a bench or jury trial. The Memphis court, on the other hand, while complying with Supreme Court mandates, prefers to follow the more traditional therapeutic concerns of the juvenile court and does not utilize plea bargaining, jury trials, or many other characteristics of the due process model followed by the Denver court.

⁵ One clear example emphasizing the different orientations of these two courts is the proportion of status offenses processed. Approximately 40% of the cases processed in Memphis in 1972 were status offenses as compared with only 10% in Denver. In the therapeutic court the legal definition of delinquency is quite broad in order to "help" all children regardless of the seriousness of their misconduct. The due process model, however, fearing the abuse of power, demands more precise rules for defining the type of conduct which may result in the finding of delinquency. Our field work and interviews with court functionaries indicate a very clear difference in attitude between agents of the two courts with respect to the type of behavior the court should focus attention upon. Hence, it is probable that the difference in the proportion of status offenses processed in these two courts reflect a difference in orientation rather than a difference in the behavior of juveniles within these jurisdictions.

³ There were approximately 9,700 eligible cases (males, not dependency, neglect, or traffic cases) from these two courts before cases with incomplete information on any of the variables were excluded. A comparison of those cases with missing data to cases without revealed no systematic differences between the two groups in the marginals of all variables. Under our inspection it appears that incomplete cases are largely attributable to random errors of record keeping.

have previously indicated, such differences in orientation are thought by many to contribute to differential treatment of suspected offenders.

In addition to orientation differences, these courts are located in different regions of the country. Thus, we are comparing a nonsouthern court with a due process orientation with a southern court with a more traditional approach to the treatment of juvenile offenders. Previous empirical research conducted in southern courts, both adult and juvenile, has concluded that racial bias was present in the sentencing of apprehended offenders (Johnson, 1941; Garfinkel, 1949; Bullock, 1961; Arnold, 1971). If philosophical and/or regional differences between courts account for differential handling of offenders, they should be readily detectable in our data.

The case records from these two courts permit us to consider the influence on juvenile dispositions of such extra-legal factors as race and social class, and the legal factors of prior record and type of offense. In addition, we have information on a variable whose influence has not been previously studied. In this study we shall label this variable as "Present Activity." The variable designates whether the child is active (working and/or in school), or is idle—that is, not working or in school. Certain observers (Platt, 1969; Cicourel, 1968) have noted the possible relevance of variables of this type to the issue of stereotyping and middle-class bias among juvenile court officials.⁶

Race and Prior Record are dichotomous variables contrasting whites vs. non-white youths and those with no prior record vs. those with prior record.⁷ Social Class is based on parental yearly income, and is composed of three categories: (1) less than \$5,000 yearly income; (2) \$5,000 to \$9,999; and (3) \$10,000 or more. Five

categories of offense type will be employed: (1) Miscellaneous, (2) Status, (3) Alcohol or Drug, (4) Property, and (5) Violent.⁸ The effects of these variables and of differences due to court orientations on severity of disposition will be assessed. The dependent variable, Severity of Disposition, has three categories which may be ranked from least to most severe.⁹ The least severe disposition is informal adjustment; the most severe is incarceration in a juvenile institution or case waived to an adult court for adjudication. Intermediate in level of severity is formal probation. In this paper we refer to these categories as "least," "moderately," and "most severe" dispositions, respectively.

Methods

Although the focus of research on juvenile dispositions primarily has been on the issue of race and class bias, we shall not restrict our attention solely to the question of whether the hypothesis of race and class bias can be supported. Instead, we shall direct our analysis to the more general question of how the six factors of race, class, present activity, prior record, offense type and court orientation appear to influence the severity of disposition accorded adjudicated juveniles. We maintain a principal focus on race and class bias, but at the same time we believe that the analysis of data such as these can offer more general information about the role of

⁸ The categorization of type or seriousness of offense was determined by grouping all possible offenses for which youths may be referred in each court into a smaller number of generic categories, and then having probation officers and juvenile judges charged with decision making at both courts rank these categories from least to most severe. There was nearly unanimous agreement among raters both within and between courts on the order of ranking of offense severity. The rankings and the various acts which comprise each category from least to most severe are presented in Appendix A.

⁹ The measure of severity of final disposition was compiled by asking probation officers and juvenile court judges from both courts to rank all possible disposition alternatives into different levels of severity. The alternatives available in both courts, and the subjective estimates as to their severity were quite similar. The categories and their proportional frequency of occurrence for both courts are presented in the Appendix (Table B).

⁶ An extensive preliminary analysis of these data indicated that certain variables available for use in the analysis had no substantial direct or indirect effect on the severity of accorded disposition in either court. These variables (age of the juvenile, family intactness, type of agency or individual making the court referral) were thus excluded from the present analysis.

⁷ Individuals of Spanish heritage are classified as nonwhite in this study.

legal and extra-legal factors in the disposition process.

To examine these data we employ Goodman's (1972) framework for log linear analysis. Log linear is well suited for the analysis of data pertaining to issues of criminal justice (Burke and Turk, 1975). Specifically, log linear analysis provides a means for examining the relationships among sets of qualitative (categorical) variables. Since many dependent variables of interest to criminologists cannot be intervally scaled properly, regression analysis and other forms of the general linear model cannot be readily applied.

Two different approaches to the analysis of the association among a set of variables can be taken within Goodman's (1972) framework. The confirmatory approach calls for the a priori specification of hypothesized models to account for the association among a set of variables. A second approach is exploratory, and involves examining a series of different models with the intent of finding the best fitting one. We find the second approach to be more consistent with the aim of answering the general question of how the six factors measured in this study influence dispositions. Furthermore, this exploratory process provides the same information for the assessment of possible race and class bias which would be obtained if we were to use the more restrictive confirmatory approach, focusing exclusively on hypotheses pertaining to the issue of bias.

Consequently, in our analysis we will examine the general impact of the six legal and extra-legal factors on dispositions. Following this analysis we will assess the specific implications of these general findings for the issue of bias in the treatment of juvenile offenders.

Goodman's (1972) framework includes three procedures which we employ in examining the influence of our set of legal and extra-legal factors on severity of disposition. First, we use a procedure for testing the fit of hierarchical models with the aim of finding a model which most parsimoniously accounts for the association among this set of variables. Second, we apply a procedure for assessing the independent (net) contribution of each

factor to the total association. Third, we present estimated effect parameters which provide information on both the sign of the relationship between variables and the relative changes of being accorded different types of dispositions.

Analysis

The starting point for our analysis was the construction of a seven-way contingency table, consisting of 720 cells, for the variables of Race (R), Parental Income (I), Present Activity (A), Offense Type (O), Prior Record (P), Court (C), and Severity of Disposition (D). Under the most general or saturated model (Goodman, 1972) the cell frequencies from this table can be expressed as:

$$G_{ijklmno} = \theta + \lambda_i^R + \lambda_j^I + \lambda_k^A + \lambda_l^O + \lambda_m^P + \lambda_n^C + \lambda_o^D$$

$$+ \lambda_{ij}^{RI} + \dots + \lambda_{no}^{CD}$$

(Total: 21 Two Variable Effects)

$$+ \lambda_{ijk}^{RIA} + \dots + \lambda_{mno}^{POD}$$

(Total: 35 Three-way Interactions)

$$+ \lambda_{ijkl}^{RIO} + \dots + \lambda_{jlmno}^{OPCD}$$

(Total: 35 Four-Way Interactions)

$$+ \lambda_{ijklm}^{RIOA} + \dots + \lambda_{klmno}^{AOPCD}$$

(Total: 21 Five-Way Interactions)

$$+ \lambda_{ijklmn}^{RIOAP} + \dots + \lambda_{jklmno}^{IAOPCD}$$

(Total: 7 Six-Way Interactions)

where λ equals the log linear effect parameter (log odds ratio). The superscripts refer to the variables and the subscripts refer to the categories of the variables. The values of $G_{ijklmno}$ are the log of the frequencies in each of the 720 cells of the seven-way cross tabulation.¹⁰ Single superscripted λ 's correspond to the effect of the marginals of each variable. Double superscripted λ 's represent the two variable or main effects.¹¹ Triple or higher superscripted λ 's designate interaction ef-

¹⁰ θ is a constant added to insure that the cell entries sum to the total sample size (Goodman, 1972: 1043).

¹¹ Consistent with analysis of variance terminology, throughout this paper we use the term "main effect" to label what Goodman calls a two variable interaction in his framework.

Table 1. Likelihood Ratio χ^2 Values for Selected Models Pertaining to the Association Among Race (R), Parental Income (I), Court (C), Present Activity (A), Prior Record (P), Type of Offense (O) and Severity of Disposition (D)

Model	Fitted Marginals	Degrees of Freedom	Likelihood Ratio χ^2	p
1	All 4-Way Interactions	284	165.42	>.5
2	All 3-Way Interactions	501	418.50	>.5
3	All 2-Way Interactions	649	1557.94	<.001
4	Model 3 + [IAR] [IRC] [APD] [AOR] [POC] [ROC] [IRC] [OPD] [OCD]	607	641.28	>.163

fects of the order corresponding to the number of superscripts.

The terms of this model involving severity of disposition (those including D) give the partial associations among the six factors and disposition. Our principal interest is in these terms. The virtue of the saturated model lies in the ability to examine both the main effects of each factor and all the possible ways in which these factors may interact to influence disposition. However, the large number of terms involving severity of disposition makes the presentation and interpretation of the estimated parameters for these terms from the saturated model unwieldy. To simplify matters we use a procedure, described below, for testing the fit of progressively simpler hierarchical models to find a more parsimonious model which adequately accounts for the association among the variables in our seven-way contingency table.

The likelihood ratio χ^2 values for tests of the goodness of fit of selected models examined in our search procedure are presented in Table 1. We began our search by examining the statistical significance of the effect parameters estimated under the saturated model. This inspection revealed that a model excluding all five, six, and seven-way interactions would adequately fit (Model 1). The likelihood ratio χ^2 shows that we clearly cannot reject the null hypothesis that Model 1 fits the observed cell frequencies. We then attempted to achieve further simplification by using a step down procedure, deleting one order of effects with each step. Model 2 is a model with all four-way interactions and above deleted. Model 3 has all three-way and higher order interactions

excluded.¹² Model 2 fits the data while Model 3 does not, indicating that the most parsimonious model includes at least one three-way interaction.

Since many of the 35 three-way interactions included in Model 2 may not be statistically significant, it is possible to achieve further simplification by finding a model which includes a smaller subset of the three-way interactions. We examined several such models and selected Model 4 as the most parsimonious model incorporating three-way interactions.¹³ Model 4 specified that we can best understand the effect of the six factors of race, social class, present activity, prior record, offense type, and court on severity of disposition by considering only their six main effects and the three interaction effects of present activity by prior record [APD], prior record by offense type [OPD], and offense type by court [OCD].

As a second step we apply Goodman's procedure for assessing the relative magnitude of effects. This procedure calls for the estimation of models deleting the term

¹² Note that because the models estimated are hierarchical, if any interaction of a set of factors is included in a model, lower order interactions consisting of subsets of these factors must also be included.

¹³ To obtain the final model we began with an examination of the effect parameters for Model 2 to see which terms among the three-way interactions might be reasonably deleted on the basis of their statistical significance. These terms were deleted and the χ^2 for the resultant model supported the hypothesis that this simplified model fits the data. We were left with twelve remaining terms, and we estimated several models deleting subsets of these twelve terms. The final model selection was made because the deletion of any of the nine terms in Model 4 resulted in significant increment to the χ^2 value for the model.

Table 2. Likelihood Ratio χ^2 Values for Models Used to Calculate Goodman's Coefficient of Partial Determination

Model	Fitted Marginals	Degrees of Freedom	Likelihood Ratio χ^2	p
5	Model 4—OPD	615	727.20	<.001
6	Model 4—OCD	615	766.83	<.001
7	Model 4—APD	609	692.65	<.001
8	Model 3—CD	651	1671.89	<.001
9	Model 3—RD	651	1559.08	<.001
10	Model 3—OD	657	2092.56	<.001
11	Model 3—PD	651	1742.12	<.001
12	Model 3—AD	651	1611.72	<.001
13	Model 3—ID	653	1567.66	<.001

of interest. These models (5–13) are presented in Table 2. The comparison of a χ^2 value of such a model with the χ^2 value for a model including the term of interest can be used to construct what Goodman (1972:1056–8) calls a "coefficient of partial determination." The larger the value of this coefficient, the greater the relative magnitude of an effect. The rank order of these coefficients is indicative of the relative importance of a factor.¹⁴

In Table 3 we present the values of the coefficient of partial determination for each of the terms of Model 4 which involve type of disposition. From this table we see that the two variables with the strongest net impact on the severity of disposition accorded are offense type and prior record. This observation applies for both the main and interaction effects involving these variables. Also we can see that the partial associations of race and dispositions, and income and dispositions make very small relative contributions to the total association. In fact Table 3 shows that the main effect of race on dispositions is nonsignificant at conventional levels.¹⁵

As the third and final step we examine the effect parameters (λ coefficients) for the relationships between the six factors and dispositions. To simplify presentation these coefficients are arrayed in two ta-

bles. Table 4 contains values of the λ coefficients for the main effects of each variable. In Table 5 we present λ coefficients for the three-way interactions involving severity of disposition specified in Model 4.

The coefficients (λ 's) found in these tables reflect comparisons with the average or overall probability of individuals being accorded the three types of dispositions. A positive coefficient indicates that, controlling for all other independent variables, there is greater than average chance that individuals falling in the designated category of a variable will be accorded the specified disposition. A negative coefficient indicates a less than average chance, and a coefficient equal to zero shows no difference from the average.

From Table 4 we again see that race has virtually no independent effect on dispositions, and that income has a significant but small effect. Further, it should be noted that the relationship between income and type of disposition is in the opposite direction of that suggested by the conventional view of class bias. Controlling for all other factors, higher income youth run a slightly less than average chance of being accorded the least severe disposition ($\lambda = -.118$), while lower income juveniles experience a slightly higher than average chance of receiving favorable treatment.

The relationships between prior record, present activity, offense type, court and dispositions all involve three-way interactions. A three-way interaction (e.g., APD) means that the relationship between two variables differs across levels of the third variable. The λ coefficients for three-way interactions specify the difference be-

¹⁴ We speak of relative importance of a factor only in the sense of its direct effect controlling for all other factors. As Duncan (1970) notes, to assess relative importance of a factor in terms of its total impact one must also consider indirect effects via an hypothesized model.

¹⁵ The test for the significance of an individual partial term is based on the value of the likelihood ratio χ^2 generated by the model comparisons specified in Table 2.

Table 3. Values of the Coefficient of Partial Determination and Likelihood Ratio χ^2 for Factors Affecting Severity of Dispositions

Factor	Comparison	Degrees of Freedom	Likelihood Ratio χ^2	Coefficient of Partial Determination	p
[OPD]	Model 5—Model 4	8	85.92	.119	<.001
[OCD]	Model 6—Model 4	8	125.55	.180	<.001
[APD]	Model 7—Model 4	2	51.37	.077	<.001
[CD]	Model 8—Model 3	2	113.95	.070	<.001
[RD]	Model 9—Model 3	2	1.11	.002	>.500
[OD]	Model 10—Model 3	8	534.62	.327	<.001
[PD]	Model 11—Model 3	2	184.18	.115	<.001
[AD]	Model 12—Model 3	2	53.78	.031	<.001
[ID]	Model 13—Model 3	4	9.72	.006	<.050

tween the main effect of a variable and its effect within the different categories of a third variable. For example, the coefficient for the interaction of the categories no prior record and active specifies that we add the value .198 (from Table 5) to the main effect parameter for the category of no prior record on the relative chance of being accorded an unofficial disposition ($\lambda = .136$, from Table 4). The resultant value (.334) gives the impact of being in the

category no prior record on the relative chance of obtaining an unofficial disposition among juveniles who are active.

Inspecting the main effects of present activity [AD] and prior record [PD], and the three-way interaction [APD], we obtain the following information about how these two variables influence dispositions. First, juveniles who are idle have a greater than average probability of being accorded the most severe disposition ($\lambda =$

Table 4. Effect of Parameters (λ 's) for the Severity of Disposition*

Main Effects	Types of Disposition		
	Least Severe	Moderately Severe	Most Severe
Race [RD]			
White	.041	-.001	-.040
Nonwhite	-.041	.001	.040
Parental Income [ID]			
Low	.088	-.031	-.057
Medium	.031	-.015	-.016
High	-.118	.045	.073
Offense Type [OD]			
Alcohol & Drugs	.096	.060	-.156
Status	.296	-.126	-.070
Property	-.195	.385	-.190
Violent	-.448	-.079	.527
Miscellaneous	.252	-.241	-.011
Court [CD]			
Denver	.111	.072	-.183
Memphis	-.111	-.072	.183
Prior Record [PD]			
No	.136	.061	-.197
Yes	-.136	-.061	.197
Present Activity [AD]			
Active	.217	.159	-.376
Idle	-.217	-.159	.376

* Parameters for the main effects of Race [RD], Parental Income [ID], Offense Type [OD], Court [CD], Prior Record [PD], Present Activity [AD]. Least Severe=unofficial disposition; Moderately Severe=formal probation; Most Severe=institutionalization or waived to adult court.

Table 5. Effect Parameters for the Three-Way Interactions of Severity of Disposition with Present Activity by Prior Record [APD], Offense Type by Court [OCD], and Offense Type by Prior Record [OPD]

	Types of Disposition		
	Least Severe	Moderately Severe	Most Severe
[APD]			
Active, No Prior Record	+.198	-.041	-.157
Active, Prior Record	+.198	-.041	-.157
Idle, No Prior Record	-.198	+.041	+.157
Idle, Prior Record	-.198	+.041	+.157
[OCD]			
Denver, Miscellaneous	+.040	-.074	+.034
Denver, Status	-.437	-.034	+.471
Denver, Alcohol & Drug	+.251	-.148	-.103
Denver, Property	+.266	-.023	-.243
Denver, Violent	-.120	+.279	-.159
Memphis, Miscellaneous	-.040	+.074	-.034
Memphis, Status	+.437	+.034	-.471
Memphis, Alcohol & Drug	-.251	+.148	+.103
Memphis, Property	-.266	+.023	+.243
Memphis, Violent	+.120	-.279	+.159
[OPD]			
No Prior Record, Miscellaneous	-.146	+.062	+.084
No Prior Record, Status	+.267	-.191	-.076
No Prior Record, Alcohol & Drug	-.154	+.037	+.117
No Prior Record, Property	+.155	+.145	-.300
No Prior Record, Violent	-.122	+.070	+.052
Prior Record, Miscellaneous	+.146	-.062	-.084
Prior Record, Status	-.267	+.191	+.076
Prior Record, Alcohol & Drug	+.154	-.037	-.117
Prior Record, Property	-.155	-.145	+.300
Prior Record, Violent	+.122	-.070	-.052

.376 and a less than average probability of obtaining the least severe disposition ($\lambda = -.217$). Second, the λ coefficients for the main effect of prior record show that, in general, having no prior record increases the likelihood of being given the least severe disposition ($\lambda = .136$) and decreases the likelihood of being given the most severe disposition ($\lambda = -.197$).

The three-way interaction [APD] indicates that the influence of prior record differs by category of present activity. The coefficients for this three-way interaction show that juveniles who are active receive less severe dispositions than would be expected on the basis of the main effect of prior record alone. Conversely, they also show that juveniles who are idle receive more severe dispositions than would be expected on the basis of the main effect of prior record alone.

The results of this analysis show that

the variable of court affects dispositions in two ways. First, the coefficients for the main effect of court [CD] demonstrate that, controlling for all other factors, juveniles brought before the Memphis court are more likely to be accorded the most severe disposition, and less likely to be accorded the least severe disposition than are juveniles adjudicated in Denver. Second, the influence of court is also shown in its interaction with the offense category.

In general, as indicated by the main effect of offense type [OD], youths adjudicated for offenses conventionally thought to be the most serious (property and violent offenses) incur the highest risks of being given either the moderately severe or most severe dispositions. But, there is also a difference between the two courts in the effect of offense category. The coefficients for the three-way interaction

[OCD] reveal that, with the exception of status offenders, apprehended juveniles in Memphis are likely to receive a more severe disposition for the same type of offense than apprehended juveniles in Denver.

Finally, the coefficients for the three-way interaction of offense type by prior record by disposition [OPD] indicate that the effect of offense type on disposition depends upon the category of prior record. The value of the coefficients for [OPD] show that this interaction principally involves status and property offenses. They further demonstrate that, on the one hand, if an individual apprehended for a status or property offense has no prior record, he is likely to receive more lenient treatment than would be expected on the basis of the main effects of offense category and prior record. On the other hand, if an individual apprehended for a status or property offense has a prior record, he is likely to receive a more severe disposition than would be expected on the basis of the pertinent main effects.

Discussion

We turn now to the implications of our findings for the issues of race and class bias and the differences in court orientations for the dispositions given to juvenile offenders. To assess the implications for race and class bias we must first consider what criteria should be used in drawing conclusions about the existence of this phenomenon. Three criteria are suggested here:

- (1) the existence of a direct effect of race or class such that nonwhites and/or lower income groups receive more severe dispositions, controlling for legal factors such as offense and prior record;
- (2) the existence of interaction effects involving race or class and legal factors—such that lower class and/or nonwhite juveniles receive more severe dispositions for the same category of legal factors; and
- (3) the existence of indirect effects of race and/or class through other fac-

tors considered to be stereotypical in nature.

The first two of these criteria are relatively straightforward, being either explicitly or implicitly employed in prior research. The third criteria, however, requires further explication. Schur (1973), as previously noted, argues that lower-class and nonwhite juveniles run a higher actuarial risk of being labeled "delinquency-prone." If race and class have no indirect effect on dispositions, but have an indirect effect through their relationship to the chance of being labeled as delinquency prone, then one could conclude that evidence of bias has been demonstrated by the third criterion.

Our analysis has uncovered no evidence of race or class bias of the type specified in criteria (1) and (2) listed above. Race has no significant association nor interaction effect with legal factors on dispositions. Parental income does not interact with legal factors in its effect on dispositions. Although there is a direct effect of parental income on disposition, the direction of this relationship is opposite to the direction hypothesized in the usual argument regarding the effect of class on treatment. Further, our analysis has shown that the total impact of this variable is quite small.

The factor of present activity is relevant to the third criterion for the existence of race or class bias. Present activity seems best interpretable as an indicator of a stereotypical perception by a court official that the juvenile is delinquency-prone. Of particular interest in this respect is the interaction of present activity with prior record. It appears that if a juvenile is idle, he receives a different type of disposition than if he legitimately is engaged in the activities of going to school and/or working at a job.

Our analysis shows that the apparently stereotypical factor of present activity has an effect on disposition by way of its interaction with prior record. To demonstrate that an indirect effect resulting in unfavorable treatment of lower-class or nonwhite youths exists, it is also necessary to demonstrate that these groups do indeed run a higher risk of being found in

the categories of present activity and prior record which lead to the most severe dispositions. Information pertaining to this question can be found in Tables 6 and 7, which show the frequency distributions of race and parental income by present activity by prior record. From these tables we can see that nonwhites and lower-income youths do run a slightly higher risk of being in the category which incurs the most severe dispositions (idle with a prior record). Correspondingly, these groups run a lower risk of being in the category which stands the best chance of receiving a favorable disposition (active with no prior record).

Given this pattern, however, an interpretation of bias is not an unequivocal matter. From Table 6 one can see that apprehended nonwhites differ little from apprehended whites in their probability of being idle. One can see also in Table 7 that lower-income youths do differ from higher-income youths in the percent idle, but it should be noted that this difference is not as marked as the difference among income groups in the percent who have a prior record. Thus, in accounting for the disproportionate tendency of nonwhite and lower-income youths to be found in the high risk category of present activity by prior record (and to be absent from the low risk category of these same variables), differences among these groups in the percentage who have a prior record are of greater importance than the differences in present activity.

It appears, then, that the greater risk run by nonwhites and lower-income youths of being labeled as "delinquency prone" does not stem as much from differences in present activity as from dif-

ferences in prior record. Since prior record is generally accorded the status of a legal factor, one can question the categorization of the indirect effect of race and class through the interaction of present activity by prior record as a type of race or class bias in juvenile dispositions.

Earlier in this paper we raised the question as to whether differences in the approach to juvenile justice taken by different courts might influence the degree of bias shown. Contrary to suggested differences in the potential for bias between a due process and a therapeutic court, no difference was found. Although the two courts differ in their philosophical or legal orientations, this difference has no effect on race or class bias. Instead it appears to center about the overall severity of dispositions and the influence of the different offense types for which juveniles were referred to the court. In our analysis, juveniles referred in Memphis (the therapeutic court) were, on the whole, more likely to be given a severe disposition, and more likely to be given a more severe disposition for the same type of offense than juveniles in Denver (the due process oriented court). Whatever discretionary power is granted under the therapeutic model, then, seems to be manifested in a greater proportion of severe dispositions accorded, but not in the degree of bias shown towards blacks and lower-class youths.

In sum, our analysis offers little support for the argument that race or class bias directly affect the dispositions given to juveniles in the Denver or Memphis juvenile courts. Rather, we found that the disposition process is most strongly influ-

Table 6. Frequency Distributions of Race by Present Activity by Prior Record

Prior Record	Race					
	White			Nonwhite		
	Present Activity			Present Activity		
	Idle	Active	Total	Idle	Active	Total
No	96 (4.2)*	1,142 (50.1)	1,238 (54.3)	103 (2.2)	1,565 (33.9)	1,668 (36.1)
Yes	204 (9.0)	835 (36.7)	1,039 (45.7)	583 (12.6)	2,366 (51.2)	2,949 (63.8)
Total	300 (13.2)	1,977 (86.8)	2,277 (100.0)	686 (14.8)	3,931 (85.1)	4,617 (99.9)

* Percent of grand total

Table 7. Frequency Distributions of Parental Income by Present Activity by Prior Record

Prior Record	Parental Income					
	Less than \$5,000			\$5,000-\$9,999		
				\$10,000 +		
	Present Activity		Total	Present Activity		Total
	Idle	Active		Idle	Active	
No	76 (2.3) ^a	1,146 (34.2)	1,122 (36.5)	86 (3.5)	968 (39.4)	1,054 (42.9)
Yes	461 (13.7)	1,671 (49.8)	2,132 (63.5)	267 (10.9)	1,133 (46.2)	1,400 (57.1)
Total	537 (16.0)	2,817 (84.0)	3,354 (100.0)	353 (14.4)	2,101 (85.6)	2,454 (100.0)

^a Percent of grand total

enced by prior record and type of offense in these two courts with different approaches to juvenile justice and from different regions of the country.

Conclusion

Our evidence suggests that offense and prior record are the major determinants of the severity of disposition accorded in the two courts studied. On the basis of these findings are we justified in concluding that the juvenile justice system of these two jurisdictions are unbiased in all aspects of their treatment of adjudicated juveniles? The answer to this question is, of course, no. However, these findings do necessitate the qualifications of the widespread accusations that race and class discrimination permeate the juvenile justice system.

Any discussion of bias in the criminal justice system must ultimately address both the issue of the point at which bias occurs and the issue of how bias is defined. Clearly, bias may appear at stages prior to a youth's referral to the juvenile court. It is possible that race and class bias occur when police and other agents decide which juveniles to refer to the court. Our data do not permit an assessment of this possibility. These data do indicate, however, that once the youth is referred to the court, prior record and offense, not race and class, are the major determinants of severity of accorded disposition.

The conclusion that bias does or does not exist in these juvenile courts depends ultimately on one's definition of bias. The criteria generally used to define the existence of bias in previous studies are based on the assumption that prior record and seriousness of offense are appropriate criteria on which to base decisions. We have followed the same approach here. Whether or not these criteria are indeed appropriate factors which should be considered when deciding the fate of juveniles is an ideological, not an empirical matter.

Some contend that the criteria utilized to make decisions in the criminal justice system are themselves biased in favor of white middle-class persons. That is, the law is seen as the embodiment of middle-class norms and the acts which represent

infractions of these norms are associated with minority group or lower-class status. Hence, the norms and life styles of many minority and lower-class group members are seen as especially conducive to behavior which may lead to imprisonment or frequent contact with social control agents. While such speculation is plausible, these issues cannot be empirically addressed in this paper. Here, *given* a set of laws or statutes defining criminal or delinquent conduct, we examine the question of possible differential treatment of minority and majority group members.

At the outset of this paper we raised the issue of whether contradictory findings in prior research are attributable to the influence of possible court differences in juvenile justice approach, or to methodological inadequacies in these studies. On the basis of our research, we suggest that the best explanation of these findings lies in their lack of methodological rigor. Although it is possible that the courts studied in prior research are different from the two studied here and that the earlier findings are accurate, we find this argument to be unconvincing. The two courts we have studied represent nearly polar positions on the continuum of possible court differences. On the one hand, the Memphis court is subject to potentially biasing factors of both region (the South) and orientation toward juvenile justice (therapeutic model), while the Denver court is characterized by factors that are thought to offer the least potential for bias (nonsouthern region, due process orientation). The lack of demonstrated bias in the extreme cases examined here at least suggests that differences in court orientation could not have accounted for the contradictory findings of previous research.

Even though we believe that the two courts studied here give us a broader base for generalization than previous single court studies, further research is called for. In addition to sampling polar cases, future research should provide for a sample of courts from other points on the spectrum of juvenile courts and should employ data-analytic techniques (like log linear analysis) which allow for a rigorous assessment of the relationships among variables in the juvenile dispositions pro-

cess. Finally, future research should attempt to consider factors thought to be related to juvenile dispositions which are not reported in the case history records used here. Some of the possible factors absent from our data and perhaps related to these decisions could be the child's demeanor, attitudes, and apparent contribution.

REFERENCES

- Arnold, William R.
1971 "Race and ethnicity relative to other factors in juvenile court dispositions." *American Journal of Sociology* 77:211-7.
- Bullock, Henry A.
1961 "Significance of the racial factor in the length of prison sentence." *Journal of Criminal Law, Criminology, and Police Science* 52:411-7.
- Burke, Peter J. and Austin T. Turk
1975 "Factors affecting postarrest disposition: a model for analysis." *Social Problems* 22:313-21.
- Cicourel, Aaron V.
1968 *The Social Organization of Juvenile Justice*. New York: Wiley.
- Cohen, Lawrence E.
1975 "Delinquency dispositions: an empirical analysis of processing decisions in three juvenile courts." U.S. Department of Justice, National Criminal Justice Information and Statistics Service. Washington, D.C.: U.S. Government Printing Office.
- Duncan, Otis Dudley
1970 "Partials, partitions, and paths." Pp. 38-47 in E. Borgatta (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.
- Garfinkel, Harold
1949 "Research notes on inter- and intra-racial homicides." *Social Forces* 27:369-81.
- Goodman, Leo A.
1972 "A general model for the analysis of surveys." *American Journal of Sociology* 77: 1035-86.
1973 "Causal analysis of data from panel studies and other kinds of surveys." *American Journal of Sociology* 78:1135-91.
- Hirschi, Travis
1975 "Labeling theory and juvenile delinquency: an assessment of the evidence." Pp. 181-203 in W. Gove (ed.), *The Labelling of Deviance*. New York: Halsted.
- Johnson, Guy B.
1941 "The Negro and crime." *The Annals of the American Academy of Political and Social Science* 271:93-104.
- Martin, John J.
1970 *Toward a Political Definition of Delinquency*. U.S. Department of Health, Education, and Welfare. Washington, D.C.: U.S. Government Printing Office.

- Packer, Herbert L.
1966 "The courts, the police and the rest of us." *Journal of Criminal Law, Criminology and Police Science* 57:238-43.
- Platt, Anthony
1969 *The Child Savers: The Invention of Delinquency*. Chicago: University of Chicago Press.
- Schur, Edwin
1973 *Radical Non-Interventions: Rethinking the Delinquency Problem*. Englewood Cliffs: Prentice-Hall.
- Terry, Robert M.
1967 "The screening of juvenile offenders." *Journal of Criminal Law, Criminology and Police Science* 58:173-81.
- Thornberry, Terrence P.
1973 "Race, socioeconomic status and sentencing in the juvenile justice system." *Journal of Criminal Law and Criminology* 64:90-8.
- Wellford, Charles
1975 "Labeling theory and criminology: an assessment." *Social Problems* 22:332-45.

APPENDIX

Table A. Severity Ranking of Offense Categories from Least to Most Severe

Offense	Denver		Memphis	
	N	%	N	%
1. <i>Miscellaneous Offenses</i> : disturbance, malicious mischief, filthy language, loitering, other	400	(16.2)	272	(6.1)
2. <i>Status Offenses</i> : behavior or condition injurious to self or others, truancy, runaway, beyond parental control, curfew	257	(10.4)	1748	(39.5)
3. <i>Drug and Alcohol Offenses</i> : marijuana possession, use or sale of marijuana, possession or sale of narcotics, possession or sale of dangerous drugs, inhaling toxic vapors, drunkenness, possession of liquor	351	(14.2)	365	(8.2)
4. <i>Property Offenses</i> : burglary, breaking and entering, auto theft, theft, fraud, forgery, shoplifting, arson, joyriding	1198	(48.6)	1709	(38.6)
5. <i>Violent Offenses</i> : assault, aggravated assault, battery, manslaughter, murder, robbery, kidnapping, forcible rape	259	(10.5)	335	(7.6)
	2465	(100%)	4429	6894 (100%)

Table B. Proportional Frequency of Severity of Dispositions

Dispositions		Denver	Memphis
Least Severe:	Case adjusted at juvenile court by probation officer or judge. The juvenile in effect is counseled and the matter then closed. Also known as "informal adjustment."	71.0%	64.8%
Moderately Severe:	Formal supervised probation.	25.8%	27.6%
Most Severe:	Case waived to adult court for adjudication or child incarcerated in juvenile institution.	3.2%	7.6%
		2465 (100.0%)	4429 (100.0%)

FORMAL AND SUBSTANTIVE VOLUNTARISM IN THE WORK OF TALCOTT PARSONS: A THEORETICAL AND IDEOLOGICAL REINTERPRETATION*

JEFFREY C. ALEXANDER

University of California, Los Angeles

American Sociological Review 1978, Vol. 43 (April):177-198

Contemporary discussions of Parsons's thought have faltered because they are insufficiently generalized. Only after his theoretical and epistemological logic has been explored, for example, can his more specific, ideological purpose be correctly understood. Parsons's theoretical ambition has been to resolve long-standing antinomies in social thought. His theory, to the degree it succeeds, formulates two central points: the social basis of individual autonomy and the multidimensional basis of social order. These positions present the heart of Parsons's formal theory. In his analysis of historical development as differentiation—cultural, structural, and psychological—Parsons combines this formal logic with an ideological commitment to the expansion of individual freedom and conscious control. Incorporating elements of liberal, idealist, and materialist arguments, he outlines a theory of substantive voluntarism that is, potentially, neither conservative nor static. Considered as a whole, Parsons's theory contains serious contradictory strains. Nevertheless, his analysis of formal and substantive freedom represents a major contribution to social thought.

The charismatic power of a great thinker raises to a heightened pitch the normal level of irrationality produced by paradigm conflict. In defense against such a powerful intellectual center, there emerges alongside the usual thrusts of serious theoretical combat an antagonistic tradition of misinformed, often trivial, sometimes grossly distorted commentary that attempts to present itself, and is partly accepted as, critical truth. At the same time, the attractive power of this center is such that those who follow the thinker prove unable to present an objective critical evaluation of his intellectual contributions. Only with the passage of time, as the center loses its immediate power, can a perspective which is both critical and appreciative be attained and the thinker's permanent contributions to intellectual tradition be properly assessed.

We can observe this tortuous path of assimilation in the reception of Marx's work and Weber's, but the process seems particularly striking, and particularly apropos of the subject of this essay, in

the case of Durkheim. As a forceful figure both intellectually and personally, Durkheim created a powerful sociological school which followed Durkheimian theory in a manner which greatly extended its scope and application but did little to articulate its foundations or to clarify its critical weaknesses (Clark, 1973). At the same time, Durkheim's theory was subject to a barrage of what was often distorted and tendentious criticism, directed not only towards his theoretical conceptions but in addition towards his ideological involvement in the reconstruction of the French republic (Lukes, 1972). Only in the late 1930s and the 1940s, with the work of sociologists like Parsons and Merton and anthropologists like Radcliffe-Brown and Evans-Pritchard—and, indeed, after the decline of almost all literalistic Durkheimianism—was the attempt begun to reappropriate Durkheim's theoretical work. And only in the last decade has the debate about his ideological perspective been sufficiently separated from the rigid radical/conservative dichotomy to enable the true humanitarian and progressive impetus of his work to be understood (Lukes, 1972: Chaps. 17, 26; Giddens, 1971: Chap. 7; Marks, 1974; Bellah, 1973).

* Many of the ideas in this essay have germinated in conversation with Robert N. Bellah, Neil J. Smelser, and Philippe Nonet. I would like to express my gratitude and appreciation.

A remarkably similar process of assimilation and its vicissitudes appears to be the fate of Talcott Parsons's work. After an initial period of the inflation of his intellectual prestige and the creation of a large number of distinguished followers, there emerged a second period during which it suffered an intensification of the same kind of distorted critical appraisal on both the theoretical and ideological levels. There have been recent indications that a more balanced sort of critical assessment is in the process of emerging, as thinkers of different theoretical and political traditions have returned to Parsons's work and argued for the centrality and significance of its central concerns (Atkinson, 1972: 1-143; Jessop, 1972; Gintis, 1969; Rocher, 1975; Bershadsky, 1973; Turner, 1974: 15-76, 193-210; Turner and Beeghley, 1974; Lipset, 1975; Johnson, 1976; Menzies, 1977; Alexander, 1979). It is as a contribution to this theoretical and ideological reassessment that this essay is intended.

MISINTERPRETATION AND THE VOLUNTARISM PROBLEM

The sociological conventional wisdom has pegged Parsons as a functionalist, an equilibrium or consensus theorist, an ideologist. Such characterizations have been promoted not only by his critics but often by his supporters as well (Lockwood, 1956; Coser, 1956; Dahrendorf, 1958; 1959; Mills, 1959; Martindale, 1960; Gouldner, 1967; 1970; Rex, 1961; Foss, 1963; Friedrichs, 1970; Wallace, 1969; Bottomore, 1974; Levy, 1952; Smelser, 1959; Mayhew, 1968b; Rocher, 1975; Johnson, 1973; 1975; Lipset, 1975; Baum, 1976; Loubser, 1976; Van Zule Slabbert, 1976). None of these descriptions, however, is sufficiently generalized to comprehend Parsons's most fundamental theoretical contribution. This lies, I would argue, more in the realm of what might be called sociological epistemology, in the formulation of a distinctive epistemological position and in its translation into the realm of sociological explanation.¹ To fully illuminate this realm of Par-

sons's work would involve exploring a wide range of diverse issues. In the present context I will limit my focus to only one aspect of the problem; namely, to the issues of voluntarism in Parsons's thought.² After elaborating, in a condensed manner, the nature of Parsons's theoretic-epistemic position on this issue, I will demonstrate its relevance for practical sociological work by indicating how Parsons combines his commitments in this realm with the ideological and empirical positions which govern his theory of social change. I will call the position Parsons has articulated on the theoretic-epistemic level his "formal voluntarism" and its empirical-ideological articulation, his commitment to the standard of "substantive voluntarism."

In social theory, the issue of voluntarism revolves, on the most general level, around two long-standing debates, the arguments over nominalism vs. realism, and subjectivism vs. objectivism (Aristotle, 1962: Bk. 7, 3, 4; Plato, 1945: 80-5, 88-92, 321-36; Augustine, 1948: Chap. 14; Halévy, 1901-1904; Stark, 1962; Sartre,

sons's functional commitment has had an important impact on his theoretical system, the influence has been on the level of model rather than on the level of epistemic presuppositions—a level which, because of its greater generality, is theoretically more significant. Furthermore, even as a model, the functional system provides wide limits of flexibility which can in no sense be associated with propositions about empirical equilibrium or conflict, or with notions about idealism-materialism, individualism-sociologism, conservatism-radicalism—as has so often been assumed (Sztompka, 1968; 1974; Stinchcombe, 1968: Chap. 3; Smelser, 1972; Hobsbawm, 1973; Lipset, 1975). This literature indicates that the notion of system is an open one, which assumes a particular content only in relation to specific theoretic-epistemic, ideological, and empirical commitments. Little significant understanding of the varied course of Parsons's intellectual development can be derived from studying his theory's functionalist aspects, whereas a great deal of that variation can be illuminated by focusing on its theoretic-epistemic assumptions.

² Though voluntarism plays a crucial role in Parsons's work, it should not be considered the master key that unlocks his entire theory. For a fuller discussion of the different levels involved in his analysis—and a methodological justification for the distinction between the theoretic-epistemological, ideological, and empirical level of sociological discourse which I assume here—see Alexander (1979: Vol. 2, Pt. 3; and Vol. 1, Chap. 2).

¹ For example, while there is no doubt that Par-

1968; Martindale, 1960; Hughes, 1958; Habermas, 1973a; Wilson, 1970). Parsons's position on these questions has been as radically misinterpreted as his position on the less generalized issues of system models, consensus theory, and ideological commitments. Basically two kinds of charges have been made. According to one strand of the critical literature, oriented to the nominalist-realist debate, Parsons is decidedly an antivoluntarist. These critics portray Parsons's writings, particularly his later ones, as concerned only with the organic whole, as taking an antiindividualist, determinist position (Scott, 1963; Martindale, 1960; Friedrichs, 1970; Atkinson, 1972; Pope, 1973; Menzies, 1977). The other critical strand, oriented to the subjective-objective question, takes a rather contradictory position. According to these critics, Parsons is an idealist who envisions no significant constraints on individual action (Lockwood, 1956; Dahrendorf, 1959; Gouldner, 1970; Heydebrand, 1972). For critics of the first persuasion, Parsons's stand on the subjective-objective issue is irrelevant; what matters is simply that he postulates supraindividual constraint per se (Atkinson, 1972: 1-145). For critics of the second persuasion, however, it is Parsons's position on the nominal-realist, or individual-society, question which is unimportant. For them, Parsons's emphasis on norms in itself commits him to an unacceptable degree of voluntarism. On one matter, however, both of these critiques are in agreement: Parsons's failings, whatever they may be, are vitally connected to his ideological conservatism.

The great contradiction between these two critical positions should alert us to their problematic status. In the following, I will demonstrate that both of these critiques are mistaken, not only in their characterization of Parsons's theoretic-epistemic position, but in their ideological critique as well.

What neither group of critics has seen, apparently, is the synthetic nature of Parsons's theoretical intention, the manner in which a major segment of his work has been directed, from the beginning, toward bridging these fundamental theoretical dilemmas (see Devereux, 1961). Western

thought can be regarded as involving two great traditions (Ekeh, 1974): the individualist (nominalist) and collectivist (realist), each of which in turn includes both romantic (subjectivist) and rationalist (objectivist) strands. To correctly comprehend one major thrust of Parsons's intellectual project, and certainly to penetrate his most important contributions, his writing must be viewed as a vigorous dialogue with each of these traditions. We will see that in his formal, theoretic-epistemic work, Parsons sought to articulate a structure for social action that ascribed voluntarism to the influence of subjective ideal elements, which are internalized by the individual and which allow him or her autonomy vis-à-vis material constraints. In this manner, Parsons rejects the nominalist notion that freedom involves the complete lack of constraint. On the other hand, in his application of this position to concrete empirical situations, Parsons has utilized this structure to articulate a model of historical development keyed to the standard of individual control over both material and ideal constraints. Parsons's relation to these two critical traditions, in other words, attempts to achieve a dialectical kind of negation, an *aufhebung* which preserves kernels of theoretical truth while it transcends the theoretical position as a whole.

FORMAL VOLUNTARISM: THE THEORETIC-EPISTEMIC SYNTHESIS

The Individualist Tradition

The most important source for Parsons's explicit emphasis on voluntarism is the individualist strands of Enlightenment thought. These are the traditions which stand at the heart of nineteenth century liberal ideology and emphasize free will as the principal ethical criterion of freedom. In its social scientific form, this ideological point is transposed into a distinctively individualistic theoretical position, which perceives social action as initiated by, and society as resting upon, discrete individuals who are free to pursue their interest as they have defined it. Historically, in terms of nineteenth century thought, this indi-

vidualist social science was articulated by neo-Kantian and Utilitarian theory (Martindale, 1960:216-66; Halévy, 1901-1904). In contemporary terms, the individualist tradition manifests itself in the sociological schools of symbolic interactionism, exchange theory, and phenomenological and existentialist sociology, all of which consider the freedom of the individual person as the starting point for theoretical analysis.

Parsons's response to this tradition argues that its voluntarism is based upon a radical misunderstanding of the theoretical role of the concept "individual," a problem which he attributes to a metamethodological problem; namely, the empiricist confusion of concrete and analytic frames of reference (Parsons, 1937: 72-4, 87-125; see also Whitehead, 1925; Schwanenberg, 1971; 1976; Fararo, 1976; Burgher, 1977; Parsons, 1970a). The concrete individual, i.e., the living, breathing, visible person, is of course free and autonomous in a certain limited sense. This describes the picture of the individual in society described by the nominalist critics. But when such empiricism is penetrated, when this person is viewed analytically rather than concretely, we can see that he is, in fact, a composite of different social forces, the most important of which are the symbolic forces which contain normative elements. Since these elements are internalized, they are in empirical concrete terms invisible: hence, when we look at an individual person, he appears to be discrete when in fact he is interpenetrated with other individuals by virtue of shared symbolic norms. It is at this point that Parsons arrives at his great insight into the voluntary quality of action. He reasons that if no individual can actually be free of constraint in the radical sense propounded by individualist theory, then what we normally perceive as free, intentional activity must in fact involve the actor's application of an internal normative standard of judgment.

If the fact of normative internalization eliminates the pure voluntarism of the perfectly free will, it simultaneously implies voluntarism of a more limited yet nonetheless significant type: the autonomy of individuals vis-à-vis the material elements

of their situation. If what appears to be the individual expression of free will is actually the determination of a certain type of social force, namely, normative symbolism, it is at the same time a very different kind of social force than that exercised by material conditions. Parsons's point, in other words, is that in order to preserve the voluntarism of the individualist strain of social thought, any conception of material forces as the exclusive determinants of social action must be overcome. "The voluntaristic system," Parsons (1937:82) writes, "does not in the least deny an important role to conditional . . . non-normative elements." It does, however, "consider . . . them as interdependent with the normative."

The originality of this aspect of Parsons's contribution, which has been thoroughly misconstrued by interpreters within the individualist tradition, is, then, his comprehension of the essential compatibility of an emphasis on voluntary individual will with a collectivist emphasis on normative interpenetration. It was, in fact, precisely to achieve this synthesis that Parsons (1937:343-409) in *The Structure of Social Action* launched his attack on Durkheim's unsophisticated sociologism. After this critique, Parsons could fully accept Durkheim's insight that the reconciliation between individual and society could be achieved only by accepting certain elements—norms—of the idealist position, a position often associated with political conservatism. Yet contrary to conservative theory, Parsons rejected the individualist perspective while pointedly retaining a voluntaristic emphasis. Parsons criticized Durkheim for obfuscating this voluntarism, but he did so only in order to formulate more effectively Durkheim's proposition that individualism, in an ideological sense, need not imply individualism in a theoretical one (Giddens, 1972:364; Bellah, 1973a).

Contrary to conventional wisdom, it is this solution to the individualist problem, this attempt to articulate voluntary action, that constitutes one part, and by far the most conspicuous part, of Parsons's famous solution to the problem of order. Normative interpenetration is induced by two individuals sharing or internalizing a

common symbol. Symbols are invariably organized into certain kinds of patterns, that is, into nonrandom kinds of arrangements. For this reason, Parsons states that just as a discrete individual is an impossible social fact, so is nonordered social action. Symbolic interpenetration means that a certain element of order exists between individuals (Parsons, 1937: 59–60, 94–6, 314, 337, 738–9; see also Levine, 1969). Order means nonrandomness, not equilibrium.

In addition to the confusion of order and equilibrium, there have been two other kinds of criticism of Parsons's theory of order, both of which involve misinterpretations so basic as effectively to turn his theory on its head. The first type of critique equates Parsons's use of order with an emphasis on conformity or passivity. Yet to make such an equation, as the preceding argument has demonstrated, is to misconstrue the entire thrust of Parsons's work and to commit the very theoretical error which his own formulation of individual action was intended to circumvent. Parsons's point, of course, is that whether independent or passive, action always involves an internalized component. It is, therefore, perfectly consistent that such criticisms of the order theory have been associated with the nominalist, individualist tradition, which views Parsons's emphasis on supraindividual force inherently antivoluntary (Pope, 1973; Cohen et al., 1975; Atkinson, 1972:181, 213; Bendix, 1970:121).

The other type of criticism levelled at Parsons's theory of order differs radically from the preceding by contending that Parsons's formulation actually allows too much voluntarism rather than too little. This accusation usually is made by materialist critics who themselves oppose the individualistic strand of social theory. Their attack, therefore, focuses on the ideal elements in Parsons's order proposal rather than on his emphasis on supraindividual order per se. By relating action to internal normative elements, it is said, Parsons has ignored the very type of supraindividual social forces which constrain action rather than facilitate it (Lockwood, 1956; Dahrendorf, 1959; Rex, 1961:78–155; Gouldner, 1970; Burgher, 1977). In-

deed, if Parsons's theory of action was limited to the elements discussed thus far, this criticism would be a telling one, and Parsons's theory would have to be considered a variation, albeit a sophisticated and voluntarist one, of the idealist tradition along the same lines as Durkheim's work. This, however, is not the case. We will see, in fact, that Parsons's theory of formal voluntarism embraces and transcends both major variants of the collectivist tradition.

The Collectivist Tradition

In addition to incorporating the individualist theory associated with nineteenth century liberalism and the normative emphasis associated with romanticism, Parsons's synthesis is also intended to subsume theoretical strands most often associated today with certain kinds of exchange and Marxian approaches—in particular the perception of individual freedom as dependent on certain kinds of supraindividual material conditions. Yet, as in the case of individualist theory, this incorporation is partial and accompanied by a formal theoretical critique. In *The Structure of Social Action* and throughout his later work, Parsons (1937:87–125) labels the exclusive emphasis on material conditions the “Hobbesian tradition” and he devotes himself to illuminating its inherent weaknesses, not by referring to some problem of order and consensus, and certainly not by weighing its ideological merit, but rather by analyzing the failures of its sociological epistemology. Parsons develops his critique primarily through a discussion of variants of nineteenth century social Darwinism. To make our analysis more contemporary, we will take some literary license and illustrate Parsons's reasoning, in a manner in which Parsons himself does not, with reference to the Marxist version of the Hobbesian tradition. In doing so, we will further indicate the tension between ideological and theoretic-epistemic approaches to freedom.

A fundamental irony of the Marxian strand of the socialist tradition is that, while its theory of material constraints has vastly increased our insight into the requi-

sites for voluntary action, its theory of social action—at least action before the achievement of a truly Communist society—perceives actions as being externally determined, and holds that the structure of objective factors squashes the human potential for intentional action or praxis. Aside from the ideological reasons for such a position, Marx's denial of voluntarism in capitalist society has two fundamental causes which are theoretically *sui generis*. In the first place, voluntarism is impossible because Marx's theory remains implicitly committed to a concrete perception of the individual actor, which means that freedom can be portrayed only as a condition in which external, material constraint is actually abolished. In the realm of scarcity, therefore, freedom is not possible. However, the antivoluntarist implication of the Marxian theory of freedom does not derive primarily from this concrete perception of the individual *per se*. Rather, it stems more from Marx's commitment to a certain perception of that individual's motivation; namely, a rationalist one—once again, a theoretical commitment that holds only for the period of capitalist and socialist scarcity. The same tradition that emphasizes material conditions portrays human action *vis-à-vis* these forces as thoroughly utilitarian. Therefore, although Marx envisions human action as potentially voluntary, as praxis, he conceived the capitalist processes of objectification, alienation, and fetishism as reducing the scope of action purely to a focus on means to the exclusion of ends.

It is here that Parsons's second great insight in *The Structure of Social Action* comes into play. He argues that such an exclusive focus on technical rationality reduces ends to the status of means. As a result, the international element of action is eliminated and determination by conditions becomes all-pervasive (see Halévy, 1901–1904). The logic here is fundamentally theoretic-epistemic rather than ideological or empirical. It implies, in other words, that by virtue of its instrumentalist position, Marxian theory is in the same logical camp as theories of political Hobbesianism or *realpolitik*, despite the latter's ideological incompatibil-

ity with the libertarian aspirations of Marx's socialist theory. Because of its similarly rationalist perspective on motivation, *realpolitik* theory also perceives social processes as determined by conditions which are over and above the voluntaristic control of ethical norms (Aron, 1971; Mommsen, 1971).

Parsons's strategy *vis-à-vis* such collectivist instrumentalism is to transform its problematic elements while incorporating these features that are compatible with a voluntaristic position. To accomplish this task, Parsons (1937:91–4, 99–100, 106, 109, 290–1, 344–9, 508–9, 576, 658) refers back to his earlier solution of the individual and order problems and, simultaneously, moves to fill out this earlier formulation by embracing elements from the materialist tradition. Relying on this first formulation, Parsons can state that because action is symbolically guided and internally directed, it contains a normative component and therefore cannot be reduced to a reflex of external material conditions. But by also orienting himself in a positive way to the materialist tradition, Parsons completes this earlier construction by asserting that this symbolic action always occurs within a conditional environment of material facts which produces pressure for the pursuit of efficient means. The key intellectual figure in this final element of Parsons's theoretical synthesis is Weber, who outlined a method for carrying forward Marx's materialist concerns without the latter's exclusive emphasis on instrumentalism. It was for this reason that, despite his own inconsistency in applying this method, Weber constituted, along with Durkheim, not only a principal reference in *The Structure of Social Action*, but a continuing source of theoretical and empirical guidance for Parsons throughout the rest of his career. It is Weber's voice we hear, not Durkheim's, in Parsons's protest that a voluntarist theory must be resolutely anti-idealist. Whereas "the voluntaristic type of theory involves a process of interaction between normative and ideal elements," Parsons (1937:82, 466) writes, "at the idealist pole the role of conditional elements disappears . . . and 'action' becomes a process of emanation. . . ."

Parsons's (1937:77) definition of the unit act, the basic element of social life, is thoroughly multidimensional:

In a unit act there are identifiable as minimum characteristics the following: (1) an end, (2) a situation, analyzable in turn into (a) means and (b) conditions, and (3) at least one selected standard in terms of which the end is related to the situation [i.e., a norm].

The Theoretical Synthesis and Parsons's Program

Action may be described, in other words, as both instrumental and normative. In terms of the subjective-objective debate, action's voluntarist quality is preserved by the latter, its determinist quality by the former. In terms of the nominalist-realist debate, action is both individual and social. Individual action is ordered by the patterning of normative symbols and by the organization of material constraints. Yet, since normative patterns are internalized, a significant cause of any action rests with the willed behavior of the concrete individual.

In order to understand fully this attempt at theoretical synthesis, it is vital to assess its peculiar status in Parsons's work. By transforming the distinctive elements of individualist, idealist, and materialist theory into a broader whole, Parsons's intention has been to delineate the structure of action and society apart from any of its particular manifestations, in the same manner that Chomsky has focussed on generative grammar as the universal structure of language (Chomsky, 1968; Ber-shady, 1973). In Parsons's (1937:733) own terms, the elements of action he has articulated have phenomenological status, in Husserl's sense. According to this analytic perspective, voluntarism is a formal property of action; it does not depend on the particular historical nature of the ideal and material conditions which constrain it.

A major part of Parsons's intellectual effort has been devoted to developing this understanding of the formal structure of multidimensional causality and value internalization into a fully elaborated theory of social life. We have dealt here mainly with his first and classic formulation in

The Structure of Social Action. In his middle period, he developed this approach much further, in the analysis of allocation and integration, the dichotomy of instrumental and expressive action, the differentiation of cultural, social and personality systems, the notions of cultural generalization and organizational specification, the analysis of pattern variables as characterizing both cultural and organizational patterns, and in his analyses of the socialization of autonomy (Parsons and Shils, 1951:53-110; Parsons, 1951; 1954; 1955; 1964). In his later work, he continued the search for synthesis in his A-G-I-L formulation, which conceptualized society as resting on the interaction of four subsystems: economics, politics, integration and value maintenance. Despite the widespread belief to the contrary, this new vocabulary represents not so much an effort to articulate the logic of functional systems per se, but an attempt to delineate in a systematic manner the degrees of material and ideal focus in any social system. In his formulations of the concrete media representing each such analytic dimension of interchange between subsystems, Parsons carried this interrelationship to its most refined level (Parsons et al., 1953; Parsons and Smelser, 1956; Parsons, 1969:157-522; 1967:385-520; Parsons and Platt, 1973).

SUBSTANTIVE VOLUNTARISM: THE IDEOLOGICAL-EMPIRICAL SYNTHESIS

A theory of formal voluntarism is necessary but not sufficient for a theory of substantive voluntarism because such a substantive theory is linked not only to presuppositions about theoretic-epistemic strategy but also to explicit ideological standards and propositions about the empirical world. Parsons's approach to substantive voluntarism is embodied in his theory of social change as differentiation. This change theory has been widely misinterpreted. It has long been argued, of course, that Parsons does not have a theory of systemic change at all. Recently, however, in response to the voluminous scholarly writing on differentiation, the anti-Parsonian critique has evolved into

an argument that the functionalist approach to change is, in fact, overly systemic (Smith, 1973). On the other side of the debate, Parsons (1967; 1970b; 1971c:27) himself has tried to describe his approach to change as being completely nonideological, as basing its analysis of evolving social structures exclusively on the criteria of which structures bring greater adaptive capacity to the social system. Neither of these interpretations is valid. Once again, we must step outside the polarization between Parsons and his critics to gain the proper perspective.

We have seen that Parsons critically reformulates the theoretic-epistemic content of the rationalist, progressive strands of individualist and collectivist thought. By no means, however, does he simultaneously abandon their ideological commitments to the expansion of voluntarism. To the contrary, it might be argued that Parsons has reformulated these theories precisely to preserve the essential libertarian aspects of their ideological perspectives. In formulating his own ideological perspective, in other words, Parsons's relation to these traditions is very different than in his formulation of the theoretic-epistemic problem.

If formal voluntarism refers to a universal property of all action abstracted from time and space, and from any specifically ideological properties, substantive voluntarism refers exactly to the opposite: to the degree that particular historical and social conditions allow the realization of individual freedom defined in terms of a particular ideological perspective. Therefore, although Parsons has discarded the individualistic position as a formal framework, his theory of differentiation accepts it as providing the basic parameters within which any theory of substantive freedom must be rooted (see Tiryakian, 1975:27-31). In contrast to his formal theory, Parsons's (1967; 1971b; Parsons and Platt, 1973:42ff) substantive theory does, in fact, take the concrete person as the point of reference. Parsons accepts, in this case, the classical liberal emphasis on the autonomy of the concrete individual, although this autonomy is, once again, a multidimensional one. Substantive voluntarism obtains to the extent

that the concrete person exercises autonomy vis-à-vis both the normative and conditional aspects of his situation. To determine the degree of autonomy of the concrete person, however, Parsons (1967) must examine the nature of its collective constraints. Just as Parsons accepts the individualistic position in his substantive, if not his formal theory, so too he incorporates the collectivist ideological commitments to freedom into his substantive theory while rejecting its formal theoretical framework. Whereas in the formal theory, Parsons's challenge is the task of interweaving norms and conditions, in his substantive theory his concern becomes the quality of norms and the quality of particular conditions.

To comprehend this substantive strategy, it is necessary to appreciate that both the traditions of critical idealism and materialism can be seen as defining freedom as the achievement of different types of differentiation.³ Within the socialist materialist tradition, for example, the most persistent strand of Marx's

³ The normative approach to freedom is obviously complex, but its very centrality in Western thought makes it a vital part of any ideological evaluation. Freedom can, of course, be viewed as natural or given, as an inherent part of any individual action. Beyond this individualist position, freedom can be viewed as dependent either upon external, supraindividual circumstances or upon conditions internal to the individual. Hobbes (1651:Chap. 21), for example, points to external circumstances: a "freeman is he that, in those things which by his strength and wit he is able to do, is not hindered to do what he has a will to do." For those in the internalist tradition, however, it is precisely the nature of this will which is at issue, not the circumstances that hinder it. Thus, as Marcus Aurelius is reported to have said, influenced as he was by the Stoics: "It is possible to live well *even* in a palace." Or as the early Christian, John, proclaimed, "Know the truth and the truth shall make you free." Freedom, in other words, is a matter of the quality of insight and perception. Within each of these general traditions there are, of course, various more specific controversies; viz., does external freedom depend on the acquisition of individual liberty or on equality? For a broad discussion of these issues and their treatment in the history of Western thought, see Adler (1958).

As a sociological theorist, however, Parsons's major contribution to this discussion lies, like Weber's, in his provision of historical-empirical categories for normative argument. He argues, in effect, that the internal and external conditions of freedom depend on the extension of cultural, social, and psychological differentiation (see Bay, 1958).

(1875:16-37; see Jessop, 1972: 46-7) sociological theory can be viewed as identifying structural fusion as the source of the inequity and domination of capitalist society. Stated negatively, only by divesting the economic structures and its dominant class of their dedifferentiated relationship to and control over the other institutional dimensions of social life can individual freedom over the environment be attained. In a positive sense, the state, as an expression of an autonomous electorate, must be able to assert its control over the economy. In general, there should be free competition of ideas and an expansion of the range of opportunities for individual action. Rather than dependent on a type of property, law must become the expression of an independent sense of right (Marx, 1843; 1875; Smelser, 1972; Avineri, 1969: Chaps. 1, 2, 6-8; Bottomore, 1974: 72-84). Although this moral position on substantive development has become distorted in the Leninist and romantic forms of Marxist thought, both of which propose forms of dedifferentiation, it has been carried on by the social-democratic tradition of Marxism (Lichtheim, 1961: Chaps. 5-6; Bottomore, 1974:97-113). In addition to Marx, of course, there are other significant intellectual formulators of the socialist theory of freedom as increased differentiation, the most important of whom extend the conception more explicitly to the noneconomic dimensions of life. Perhaps the most important of these figures are Michels (1962) with his theory of political democracy as the competition of elites, and T.H. Marshall (1965), who emphasizes the social aspects of citizenship (see also Lipset, 1962).

The Western intellectual tradition that emphasizes freedom as differentiation on the ideal or normative level rather than on the material one is less explicit although no less significant. Its premise is that individual autonomy defined as the individual's control over his or her internal environment occurs to the degree that spiritual and ethical issues are transcendent vis-à-vis earthly concerns. This position has been articulated by such diverse traditions as the Judaic and Christian notions of divine law and the natural

law traditions of the French Enlightenment (Weber, 1952; 1954; Nelson, 1949; Becker, 1935); the Protestant notion of the sanctity of individual conscience and the legitimation of individual doubt and its secular expression in theories of democratic rights (Weber, 1958b; 1958a; Parsons, 1937:51-8; Little, 1969; Walzer, 1965; Tiryakian, 1975: 24-30); and the various conceptions of freedom tied to the autonomy of secular intellectual thought from the Greeks and the Humanists to Bacon (Voegelin, 1956; Strauss, 1953; Shils, 1972).

It is, then, out of these two conceptions of liberty, the individualist and collectivist, that Parsons has forged the principal part, and certainly the most enduring part, of his substantive theory of social change. He embraces the ideological position that the freedom of the concrete individual depends on the differentiation of both conditional and normative structures, but at the same time he articulates these substantive goals through the synthetic perspective which he used to criticize the formal aspects of these same traditions. In fact, it is possible to argue that only by transcending the formal, theoretic-epistemic problems of these arguments can certain long-standing ideological problems be overcome (Bay, 1958). For example, by ignoring the impact of collective forces, particularly the accumulation of economic and political power, individualist perspectives risk the social irrelevance of traditional liberal ideology (Parsons, 1928; 1929; see also Lipset and Ladd, 1972). On the other hand, by postulating only rationalist-utilitarian motivation, Marxist theory risks the ideological acceptance of anti-normative, nondemocratic force (Trotsky, 1938; Merleau-Ponty, 1947; Parsons, 1967:102-35). Finally, by ignoring the conditional, constraining dimensions of social life, idealist democratic ideology risks the possibility of an abstract utopianism (Parsons, 1967; Marx and Engels, 1848:61-4).

By grounding his theory of substantive voluntarism in the formal integration of these three positions, Parsons attempts to avoid such pitfalls. According to his theory of social change, personal au-

tonomy is achieved to the degree that the institutions associated with the different dimensions of society, the functional subsystems of economics, politics, integration, and value maintenance, become differentiated from one another and, in the process, develop (1) their own independent criteria for performance as expressed in institutionally separated media; and (2) the capacity to mobilize the resources of other dimensions by asserting a partial but independent regulation over them. Although these developments—which Parsons calls the growth of institutionalized individualism—are viewed as occurring within the context of the social system as a whole, they can be seen as involving differentiation of three distinctive types: cultural, structural, and psychological (Parsons, 1966: 20–9; 1971b; 1971c: 18–28).

Cultural Differentiation

In terms of the formal theory of multidimensional causality, cultural, or value patterns constitute a dimension of every social structure and, at the same time, an independent dimension subject to an independent set of causal forces. In terms of Parsons's substantive theory of social change, it is necessary to consider the development of these patterns as an independent process of differentiation with a distinctive relationship to the achievement of substantive voluntarism. The achievement by the value dimension of society of transcendent regulative power vis-à-vis more conditionally-oriented social structures, Parsons believes, is directly related to the capacity of a society, or a particular society group within it, to engage in reform and directed social change (see Durkheim, 1893).

Parsons (1966; 1961; 1963) follows Weber in viewing the most significant historical periods of cultural differentiation as the radical breaks in normative order created by religious upheaval, particularly the rise of transcendent religions that occurred in different civilizations during the first millenium and the later cultural break created by the Protestant Reformation. Bellah's (1970:20–50) article, "Religious Evolution," represents the most sophisticated formulation, couched within the

general framework of Parsons's formal theory, of the relation between stages of religious transcendence, social reform, and the achievement of substantive voluntarism (see also Eisenstadt, 1973: Chaps. 6–9). Little's (1969) highly original historical monograph, *Religion, Order, and Law*, demonstrates how the differentiation of the symbolic order is a basic prerequisite for political and legal differentiation and, in general, for the increase of individual autonomy. On a more specific level, symbolic differentiation has been dealt with in terms of the differentiation of specialized types of cultural patterns. Parsons (1961b; Parsons and Platt, 1973; Chap. 6), Eisenstadt (1969:64–7), Geertz (1973), and Barber (1971) emphasize the significant social leverage provided by the emergence of secular political morality or ideology. Parsons (1951: Chap. 6; Parsons and Platt, 1973: Chap. 6) and Barber (1952: Chaps. 2, 11) analyze the emergence of social and natural science in terms of the historical impact of this differentiation of an autonomous cultural pattern of secular rationality.

The most systematic and widely tested Parsonian framework for dealing with the relation between the development of substantive voluntarism and the emergence of autonomous cultural levels is the pattern variable scheme, especially the universalism-particularism dichotomy. In Parsons's perspective, the cultural pattern of universalism promotes critical judgment because it demands that all particular traits be evaluated according to a broader, more general set of principles. Universalism, in other words, is a form of differentiation: it creates distance between the cultural norms and the object of judgment. In a series of essays on Japan, Turkey, the nations of Western Europe, and the United States, Bellah (1970: 53–189; see also Geertz, 1971) has traced the effect of religious particularism and universalism on the possibility for achieving democratic political activism. Lipset's (1967) *The First New Nation* is the most important application of the pattern variable schema to the specific question of the impact of different kinds of Western political cultures on the possibility for structural reform.

Finally, it is important to emphasize, particularly because of the critical misunderstanding of Parsons's level of abstraction, that this Parsonian analysis of developmental changes in ideas is not an idealist, emanationist approach any more than is Freud's developmental theory of personality or Weber's sociology of religious change. Although Parsons's multidimensional theory allows isolation of the independent effect of developments in the cultural sphere, it simultaneously indicates that such changes must always be related to developments in other dimensions. The movements towards cultural differentiation and universalism are movements by social groups, who have responded to conflicts and strains which may have originated in any dimensional location by formulating new and more transcendent symbolic patterns. The interpenetration of self-interest, religious interest, and religious differentiation is nowhere more effectively demonstrated than in Eisenstadt's (1969: Chaps. 4-9; see also Eisenstadt 1964a; 1964b; 1965; 1973:119-50) discussion of the struggle between bureaucracy and church in the historical empires.

Structural Differentiation

Struggles by groups to maintain and, usually to increase their scope of action also initiate differentiation in the structural sphere, which in Parsons's terms refers not to material conditions per se but rather to institutional organization in each of the four different dimensions of the social system. What is at stake in this aspect of differentiation is the long and painful growth of the autonomy of different institutional sectors, the evolution from an historical situation in which single institutions, groups, and leaders perform, and therefore monopolize, multifunctional tasks to a more diversified structural situation in which there is more of a single function focus. Differentiation in any given dimensional sphere can be described as a never-ending process. It begins with the simple emergence of a new structure (for example, a centralized government bureaucracy) and only gradually achieves a certain level of substantive au-

tonomy, as when a central government structure becomes constitutionally democratic or achieves the ability to command economic resources for public action. Each phase of structural differentiation is linked to greater self-expression, to increased voluntarism, for a particular group, and usually to an increase in the supply of resources—economic, political, integrative, or value-oriented—for the expansion and growth of groups in certain other sectors. At the same time, any given phase will usually also involve the suppression of rights and the restriction of voluntarism for certain other groups, and for this reason any instance of structural differentiation often triggers renewed struggle.

The differentiation of the economic market system, with its power to command resources from other sectors and its development of an independent form of media in money, represents a classic example of how structural differentiation can interweave autonomy, expansion of resource production, and the partial suppression of rights. Parsons (Parsons and White, 1964; Parsons and Smelser, 1956: Chap. 5) and Barber (1974) have compared the degree of substantive voluntarism associated with a differentiated economy to the situation in preindustrial, undifferentiated situations. *Economy and Society* (Parsons and Smelser, 1956) remains the most sophisticated general statement of the relation between a differentiated economic system and other sectors of the social system. The historical processes by which successful economic differentiation and the expansion of the scope of action have been achieved, and through which successive noneconomic differentiation in turn has been produced in response to the suppression of preindustrial modes of societal integration, have been the main concern of Smelser's (1959; 1968: Chaps. 6-8) discussions of modernization. Eisenstadt's (1969: Chaps. 2, 3, 8, 11, 12; see also 1964a; 1964b; 1965) *The Political System of Empires* contains an historical and theoretical analysis both of the conditions that make economic differentiation problematic and of how economic dedifferentiation prevents the expansion of substantive voluntarism

through its dampening effect on the possibility for differentiation in other dimensional spheres.

Another major focus of the multidimensional approach to the problem of substantive voluntarism has been the process of the differentiation of political structures. Political differentiation creates the structural apparatus for a society to define self-conscious goals and in doing so to discipline and regulate the resource production of other spheres. Such capacity is achieved through the differentiation of organs for executive administration, like bureaucracies, and of organs for mobilizing support and articulating grievance, like parties and constitutions. These developments depend on, first, such internal factors as the quality of political leadership, and, second, the differentiation of external resources, like economic adaptation sufficient to produce tax revenues, the increasing breadth of integrative groupings necessary for mobilizing support, and the growth of transcendent cultural standards as references for political reform. Eisenstadt's (1969) writing on the historical empires represents the broadest empirical analysis of this kind of interrelationship, a framework extended over the full historical range of political structures in Eisenstadt's (1971b) *Political Sociology*. On a more abstract level, Parsons (1967:422-65; 1969:163-78; 1971c) has written extensively on the different kinds of dimensional inputs associated with the emergence of dictatorship, bureaucracy, democracy, and collegiality as forms of political organization. In discussing the differentiated conditions necessary for the democratic competition of political elites, Lipset (1962; Lipset and Rokkan, 1967; see also Surace, 1976) has moved beyond the rationalistic formulations of Michels (1962) and placed that central ideological issue on a different plane of analysis. Smelser (1973; 1974; see also Huntington, 1968:93-139; Rueschmeyer, 1977) has analyzed recently the political forces that maintained a condition of functional de-differentiation in the California system of higher education in the 1960s. These notions have also been pursued at length within the Parsonian tradition of political science, particularly by Almond (1956;

1960; Almond and Powell, 1966), Apter (1958; 1966; 1972), Easton (1953; 1965), Deutsch (1963; 1964), Mitchell (1958; 1972), and Johnson (1966). In her work, Keller (1963; see also Eisenstadt, 1971c) formulated the implications of differentiation theory for institutional stratification studies; she argued that the process replaces "ruling class" by "functional elite."

The other dimensions of structural differentiation and the types of substantive voluntarism which they entail have been accorded less attention to the degree they depart from the traditional concerns with economy, polity, and religion. In regard to the integrative dimension, with the exception of the phenomenon of citizenship to be discussed below, the focus has been limited to the problem of the differentiation of legal structures. In addition to the discussions by Parsons (1971c: Chap. 2) himself, the work of Little (1969) and Mayhew (1968a) traces a continuum from the first delineation of distinctive secular rights to the successive attempts at their real institutionalization. In terms of differentiation in the value dimension—excluding religious patterns, which we have already analyzed as cultural developments—the emerging autonomy of families, peer groups, schools, and scientific institutions has been discussed, respectively, by Smelser (1959), by Eisenstadt (1971a) and Parsons (1964: 155-82), by Parsons and Platt (1973), by Dreeben (1968) and Ben-David (1971). The manner in which these developments facilitate the growth of substantive voluntarism will be the subject of the analysis of psychological differentiation below.

Little theoretical or empirical work has been done on bringing these various analyses of structural differentiation together to develop a more integrated theory of multidimensional causality and a fuller notion of the ramifications of the growth of substantive voluntarism in a differentiating system. Perhaps the closest attempt to such a synthesis is Parsons's (1971c: Chaps. 2, 6; 1967: 490-520; Parsons and Platt, 1973: Chap. 4) analysis of the differentiation of the integrative dimension of the nation, the dimension he labels the "societal community." To the

degree that the societal community is differentiated, the national community becomes defined universalistically; such universalism implies that in crucial instances an egalitarian national solidarity will supercede the more particularistic definitions of national community generated by class, race, ethnicity, region, or religion. Clearly, the emergence of such a societal community is a major prerequisite for the achievement of substantive voluntarism, for the ability to control and reform the production of different dimensional resources in an egalitarian way. As such, this differentiation is dependent on very distinctive kinds of developments in the economic, political, value and normative spheres. It can be viewed, in fact, as the result of the interminable struggle by social groups in each of these different spheres for continually more effective inclusion into the national society (see Eisenstadt, 1969: 248). Obviously such a process raises the continual possibility for dedifferentiation, which can be defined as a narrowing of the definition of national community in a particularistic direction, resulting in the constraint and reduction of substantive voluntarism. It is possible to draw a direct relationship between Parsons's (1954:104-44, 298-322; 1969) discussions of such dedifferentiation in integrative structures—his essays on the sources of Fascism and political aggression in Western society and McCarthyism in American society—and the conception of status politics developed by Hofstadter (1952) and Lipset. Lipset (1967: Chaps. 7, 9; see also Pitts, 1964), in fact, has used extensively an implicit dedifferentiation notion to trace the manner in which dominant class, political, and solidary groups have skewed European social developments towards Fascism or Communism. More generally, this notion of differentiation and the societal community directly connects Parson's (1967: 385-421; 1960: 295-321; 1971c: 30-2, 92-3; Eisenstadt, 1969: 80) theory of substantive voluntarism, as I have developed that conception here, to the more empirical theory of voluntary groups and voluntary associations developed in other sociological literature.

In this section, I have demonstrated

that arguing within the general multidimensional framework of his formal theory, Parsons has attempted to develop an empirical and historical theory of the structural requisites for freedom and of the structural contradictions that constrain its achievement. As such, despite the still elementary level of its articulation, Parsons's theory of historical development presents the possibility for advancing well beyond the individualist, idealist, or materialist approaches to the problem of the institutionalization of individual freedom in social life.

Psychological Differentiation

As I mentioned earlier, the most important explicit source of voluntarism as a hegemonic ideological principle in Parsons's theory has been the individualist focus of liberal thought. Parsons's dialectical relation to this tradition, theoretically and ideologically, reaches its apogee in his analysis of the psychological aspects of the differentiation process. Through his conception of what he called the Resource Chart, Parsons (Parsons and Smelser, 1956:139; see also Smelser, 1959; 1962) integrates Freudian and Piagetian theories of individual development with his analyses of cultural and structural differentiation. The Resource Chart conceptualizes the sequential stages involved in the production of an individual from the earliest stage of childhood to achievement of the adult role. In terms of Parsons's later theory of subsystem interchange, this process can be described as the passage from full-time participation in the institutions of the value maintenance dimension to participation in the institutions of the more conditionally-oriented organizations of adult society. Since the effect of cultural and structural differentiation is to separate value maintenance institutions both from one another and from institutions in other dimensions of the social system, the impact of such differentiation of the Resource Chart can be visualized as increasing the number of developmental stages involved in the passage from childhood to adulthood. In terms of Freudian theory, the greater mobility required for this transition to adulthood can be accom-

plished only by increased ego autonomy and control over affective dependency. Conversely, in terms of differentiation theory, this social demand for psychological voluntarism meshes with the psychological opportunities provided by certain structural developments: the growing separation between the increasingly functionally-specific nuclear family and the institutions that develop to fulfill other social functions facilitates rebellion, separation, and neutralization vis-à-vis basic object relationships.

This interweaving of the Freudian understanding of individual personality growth with the broader theory of substantive voluntarism as the product of cultural and structural differentiation has been developed most by Weinstein and Platt (1969). In *The Wish To Be Free*, they contend that it was the development of structural differentiation—in Parsons's technical sense, the expansion of the Resource Chart—which eventually created in nineteenth century Western society the opportunity for successful Oedipal rebellion and separation from authority. In turn, they argue that this expansion of the developmental process and increase in psychological autonomy were themselves crucial and independent variables in the subsequent development of substantive voluntarism through the differentiation of other institutional spheres. At the same time, Weinstein and Platt (1969) emphasize that just as social and psychological differentiation are related in a positive manner, so can they be negatively interrelated. The lengthening of the passage from full-time participation in value maintenance institutions to more participation in conditional dimensions is not only a liberating but also a perilous development. By radically increasing the challenge of the transitional process, it also increases the likelihood for failures, for the pathological dedifferentiation that Weinstein and Platt (1969: Chap. 7) call the universal reactions to modernization. In his essay on youth culture, Parsons (1964: 155–82; see also Slater, 1961a; 1961b; Chodorow, 1974; 1978) has described one stage of this dedifferentiation as a basic structural problem of modern society.

This kind of historical analysis of the interplay between psychic and social differentiation has also been pursued by Bellah (1970:76–99) in a suggestive essay, "Father and Son in Confucianism and Christianity." Bellah (1970) follows Erikson (1950) in describing how universalistic developments in the religious sphere provide a point of leverage for the development of greater psychological control over primary object relations, particularly over those objects associated with authority and domination.

The moral development side of the Resource Chart concerns the issue of socialization rather than personality development per se, and rests upon two basic insights by Parsons (1964:17–33, 78–111): (1) no aspect of affective development, even within the nuclear family itself, is ever separated from the development of symbolic-moral matters; and (2) structural differentiation and the increasing isolation of the nuclear family create the necessity for the differentiation of a range of transitional structures to mediate the passage from childhood to adulthood. On the basis of these propositions, Parsons (Parsons and Bales, 1955:119–23, 136, 155; see also Menzies, 1977:104–9) integrates his analysis of personality development with the pattern variable schema of cultural differentiation and Piaget's theory of moral and cognitive development. According to Parsons, the relation between mother and child transmits to the child a thoroughly particularistic moral pattern, in that it presents, from the child's perspective, the very prototype of a relationship in which no higher differentiated standard or judgment is possible. Furthermore, even though relative to this initial relationship moral development in the family becomes increasingly universalistic, family morality remains particularistic relative to other institutions because of its primarily affective function. It is clear, in other words, that if substantive voluntarism is to be achieved—in pattern variable terms, if the universalistic pattern is to be internalized—further stages of affective and moral socialization beyond the nuclear family are a necessity. The school is perhaps the most significant

institution which becomes differentiated in response to this need, and the most important discussions of schooling in terms of the pattern variables are Dreeben's (1968) *On What Is Learned in School* and Parsons and Platt's (1973: Chaps. 2, 4) *The American University*. The successful contribution to this extended socialization process of the other newly differentiated structure, the adolescent peer group, has been analyzed by Eisenstadt (1971a; see also Fass, 1977) in *From Generation to Generation*; the more particularistic pathological aspects of peer group culture by Pitts (1964:49-59) and Parsons (1951:286-93).

As this brief section indicates, Parsons's work contains a complex theory of psychological differentiation which is thoroughly interwoven with his theories of differentiation on the cultural and structural levels (see also Inkeles, 1971). It should also be clear that, within the general framework of his formal theory, this analysis of psychological developments is keyed to the evaluative conception of freedom as the achievement of individual autonomy and control. The view of Parsons's theory as an oversocialized conception of man completely ignores this vital aspect of Parsons's substantive work and, in turn, the distinctive formal apparatus upon which it is built (Wrong, 1961; Gouldner, 1970: 218-39).

Differentiation as a Conflict Theory

Parsons has utilized his theoretical framework of formal voluntarism to explore the requisites for a theory of individual control in the cultural, structural, and psychological spheres. No matter how differentiated, of course, individual action remains connected and disciplined by the environments of nature, on the one hand, and by what Parsons (1966; Chap. 1) calls the ultimate reality of existential meaning on the other. Parsons's theory of institutional differentiation is not intended to deny this fundamental connection. Nevertheless, he has clearly proposed a change theory which is geared to the problem of the expansion of individual freedom and choice. It is far from the theory of

conservation—of the status quo, of the seamless web of social intercourse, of the supraindividual system—that his critics have charged.

To the contrary, Parsons's (1961:344; Parsons and Shils, 1951:216) differentiation theory proposes, in principle, that the process of differentiation, and the increasing independence of structures and individuals which results from it, will increase the general level of social conflict, although it may at the same time increase institutional flexibility in handling and channeling conflict's repercussions. Since societies, particularly modern and modernizing ones, are continually subject to strains at all levels and at varying intensities, social change must be viewed as a constant and highly uneven process (Smelser, 1971: 7). According to the Parsonian theory of change outlined above, societies have two options in response to such strain: either differentiation or de-differentiation. Because it allows flexibility in the face of vested interests, both material and ideal, differentiation is linked to the capacity for system reform and to the extension of individual freedom. If a society is unable to engage in differentiation, the social response will be to suppress the reaction to strain rather than to eliminate its source. The possibility will be raised for the conflation of differentiated structures and the reduction, rather than expansion, of the possibilities for individual autonomy and control (Smelser, 1962; 1974; Eisenstadt, 1964c; 1969; Weinstein and Platt, 1969). By creating an historical theory that describes a continuum of different kinds of reaction to strain and the conditions under which each might occur, Parsons has developed a perspective that, far from neglecting conflict, presents the framework for a comprehensive approach to its explanation and evaluation (see also Rueschmeyer, 1977).⁴

⁴ In fact, Parsons's theory of social change can be seen as providing a more general framework for analyzing the very theoretical points advanced by the three conflict theorists who have been some of his sharpest critics (see Atkinson, 1972); the ruling class situation described by Rex (1961); the superimposition/pluralization theory proposed by

IRRESOLUTION IN THE THEORETICAL AND IDEOLOGICAL SYNTHESIS

As I noted at the outset of this essay, my purpose has been to initiate a new reading of Parsons's work. In supporting my proposal that the notions of formal and substantive voluntarism provide one element of such a framework, I have ignored aspects of Parsons's writing which are contrary to this major thrust. Although a lengthy analysis of these contradictory strands does not have a place here, I will mention the kinds of problems I have in mind (see Alexander, 1979).

The problematic aspects of Parsons's theory are elements that reflect an ambivalence about the very theoretical and ideological tensions that Parsons has resolved so effectively in the strands of his work analyzed above. Because of this ambivalence, his work, considered as a whole, contains major contradictions on both formal and substantive levels.

On the formal level, the synthetic approach that Parsons so carefully develops is crosscut by a significant idealist strain. Simply in terms of his focus of attention, the internal dimensions of values and norms, both as institutional spheres and as cultural patterns, have received vastly more of his attention than the conditional dimensions of economics and politics. More importantly, when these latter dimensions do become the focus of Parsons's extended attention, they are described empirically in such a manner that they facilitate rather than conflict with the normative inputs to their production (see Gouldner, 1970:286-325). Furthermore, the systemic conflict that Parsons does find throughout social life much more

often occurs in the value or normative dimensions and in the tension between these dimensions and the more conditional ones, than within the economic and political spheres themselves. In addition to these problems, a strain of holism characteristic of an unreconstructed realism coexists alongside Parsons's focus on the independent role of the individual in action. For example, Parsons emphasizes conflict between internally integrated subsystems and between groups which share the same overarching value commitments rather than conflict between groups within the same subsystem which, to use Evans-Pritchard's (1953) term, "refract" common value patterns into partial and opposed commitments.

The same kinds of ambiguities of resolution represent significant strains in Parsons's ideologically related approach to social change. For example, in his description of cultural differentiation as generalization, there is a tension between a rather conservative emphasis on generalization as simply the provision of greater integration and the other emphasis, described above, on its promotion of increased critical activism and, indirectly, social conflict (Parsons, 1971b; Toby, 1975). Or again, there are clear signs that Parsons has only incompletely resolved his relation to the individualist emphasis of traditional liberal ideology. In comparison to some other analysts of differentiation at the psychological level, Parsons's recognition of the alienative psychic costs of substantive voluntarism has been strikingly inconsistent. While in some essays he traces with great subtlety the alienation attendant on the realization of autonomy and differentiation (1954: 89-103, 177-96, 298-322; 1964: 112-26, 257-96; 1971a), in others he portrays the achievement of individualization and affective autonomy as relatively unproblematic (1967:3-34; Parsons and White, 1969). And although his basic commitment to the collectivist tradition is clear in his support of welfare state *laissez-faire* ideology, Parsons (1954:386-439) underplays, in a manner characteristic of traditional liberal individualism, the social costs of economic systems that institutionalize private property (1954; see Gouldner, 1970:302-4,

Dahrendorf (1959); and the notion propounded by Coser (1956) that conflict can have a positive function for societies with sufficiently flexible social structures. Compare, for example, the following quotation from Coser (1956:154) with the differentiation theory outlined above:

A flexible society benefits from conflict because such behavior, by helping to create and modify norms, assures its continuance under changed conditions. Such mechanisms for readjustment of norms is hardly available to rigid systems: by suppressing conflict, the latter smother a useful warning signal, thereby maximizing the danger of catastrophic breakdown.

320-3). Despite the effective argument that can be made that democratic socialism represents an advance towards social differentiation and towards the achievement of substantive voluntarism, Parsons has always dismissed out of hand the developmental advantages to be gained from institutionalizing public ownership and redistributive public policies (see also Rocher, 1975:144).

It should be clear, in light of the entire preceding analysis, that none of these emphases is a logical or necessary part of Parsonian theory, in either its formal or substantive versions. It is not accidental, in this regard, that our references in the preceding section were to Parsons's students and coworkers as much as to Parsons himself. It has often been Parsons's students, not Parsons, who have explored the full range of the theory's application—in the formal realm to the problems of political and economic conflict, in the substantive theory to the application of the social criticism inherent in its central logic. Within the framework of the formal theory, for example, Smelser (1973:390-7; 1971:8-9; see also Rueschmeyer, 1977) has recently criticized Parsons's underemphasis on the problem of power; and in his own work, Smelser has portrayed the differentiation of the political, not the cultural dimension as the crucial factor in the development of substantive voluntarism. Similarly, Eisenstadt (1969) has written at great length, in the formal framework of multidimensionality, about the often insurmountable conditional problems presented by economic classes and by the centralization of political power. And Lipset's (1967; see also Pitts, 1964) utilization of the pattern variable scheme contains an extensive analysis of political conflict in terms of subgroups representing opposing value patterns. On the ideological side, the last decade has produced a distinctly leftward movement among some Parsonians, who are responsible for a series of essays which critique contemporary Western society from the general perspective established by differentiation theory (Pitts, 1974; Bellah, 1970: 193-257; 1975; Eisenstadt, 1973: 231-57; Gould, 1976; Smelser, 1975).

CONCLUSION

I have argued here against a number of standard interpretations in the rapidly growing commentarial literature on Parsons's theory. In opposition to these charges, I have proposed that one of Parsons's major contributions can be most effectively appreciated in terms of the dichotomous classification of a concept which has been central to Western thought, the concept of voluntarism. In his formal framework, Parsons articulated a self-conscious integration of individualist, idealist, and materialist theories which described the properties of action in terms of an interweaving of the voluntarism produced by the pursuit of normative ideals and the constraint induced by the chains of material necessity. By developing within this formal approach a theory of historical change as differentiation, Parsons proposed a framework which potentially integrates the emancipatory aspects of individualist, idealist and materialist ideology and, in doing so, provides the basis for evaluating history in terms of its realization of voluntarism in a substantive sense.⁵ Yet despite the enor-

⁵ Because of this synthetic intention, I would argue, Parsons's change theory has potentially universal application, far beyond the range of Parsons's own usage. It is revealing in this respect to compare differentiation theory with the approach to change taken by Jürgen Habermas, the Frankfurt school Marxist. The purpose Habermas has set for himself—to preserve Marx's ideological commitment to freedom while transforming the instrumentalism of his theoretical apparatus—leads him ineluctably to a change theory that resembles Parsons's own. Differentiation theory, however, can clearly be viewed as subsuming Habermas's theory of communication distortion. In working out the latter idea, Habermas's (1973a:315) intention is to construct a theory of human "evolution toward autonomy and responsibility" keyed to the ideal of increased "freedom from domination." In order to do so, he realizes that his theory must measure progress toward "human adulthood" on the psychological level, and he incorporates Freudian concepts to accomplish this (1970:119; 1973b:256). Habermas acknowledges further that, in addition to including the structural emphasis of Marx, he must address the problem of the historical development of moral systems (1973b:2-3) and the preconditions of an autonomous public opinion, one with the capacity to mediate between a society and its social values (1970:72-4). Despite the often brilliant texture of his argument, however, Habermas has failed throughout

mous accomplishments of Parsons and the members of his sociological school, the theories of both formal and substantive voluntarism remain relatively undeveloped. Not only are vast theoretical, ideological, and empirical issues barely articulated, but Parsons's own contributions have been marred by contradictory strains. Parsons's fundamental contributions to social thought have only begun to be reappropriated.

REFERENCES

- Adler, Mortimer J.
1958 *The Idea of Freedom*. New York: Doubleday.
- Alexander, Jeffrey C.
1979 *Theoretical Logic in Sociology*. 2 Vols. Berkeley: University of California Press. In press.
- Almond, Gabriel S.
1956 "Comparative political systems." *Journal of Politics* 18:391-409.
1960 "A functional approach to comparative politics." Pp. 5-64 in Gabriel Almond and J. S. Coleman (eds.), *The Politics of Developing Areas*. Princeton: Princeton University Press.
- Almond, Gabriel, and G. Bingham Powell, Jr.
1966 *Comparative Politics*. Boston: Little, Brown.
- Apter, David E.
1958 "A comparative method for the study of politics." *American Journal of Sociology* 64:221-37.
1966 *The Politics of Modernization*. Chicago: University of Chicago Press.
1972 *Ghana in Transition*. Princeton: Princeton University Press.
- Aristotle
1962 *Politics*. Edited by Ernest Barker. New York: Oxford University Press.
- Aron, Raymond
1971 "Max Weber and power politics." Pp. 83-100 in Otto Stammer (ed.), *Max Weber and Sociology Today*. New York: Harper and Row.
- Atkinson, Dick
1972 *Orthodox Census and Radical Alternative*. New York: Basic Books.
- Augustine, Aurelius
1948 *The City of God*. Edited by Didos Marcus. New York: Hafner.
- Avineri, Shlomo
1969 *The Social and Political Thought of Karl Marx*. Cambridge, Ma.: Cambridge University Press.
- Barber, Bernard
1952 *Science and the Social Order*. New York: Free Press.
1971 "Function, variability, and change in ideological systems." Pp. 224-62 in Alex Inkeles and Bernard Barber (eds.), *Stability and Change*. New York: Little, Brown.
1974 "The absolutization of the market." Paper given at the Conference on Markets and Morals.
- Baum, Rainer C.
1976 "Introduction." Pp. 448-69 in J. J. Loubser, Rainer C. Baum, Andrew Effrat and Victor Meyer Lidz (eds.), *Explorations in General Theory in Social Science*. New York: Free Press.
- Bay, Christian
[1958] *The Structure of Freedom*. Stanford: Stanford University Press.
1970
- Becker, Carl L.
1935 *The Heavenly City of the Eighteenth Century Philosophers*. New Haven: Yale University Press.
- Ben-David, Joseph
1971 *The Scientist's Role in Society*. Englewood Cliffs: Prentice-Hall.
- Bellah, Robert N.
1970 *Beyond Belief*. New York: Harper and Row.
1973a "Introduction." Pp. ix-iv in Robert N. Bellah (ed.), *Emile Durkheim on Morality and Society*. Chicago: University of Chicago Press.
1973b "The five civil religions of Italy." Unpublished manuscript.
1975 *The Broken Covenant*. New York: Seabury.
- Bendix, Reinhard
1970 *Embattled Reason*. Oxford: Oxford University Press.
- Bershady, Harold J.
1973 *Ideology and Social Knowledge*. New York: Wiley.
- Bottomore, T. B.
1974 *Sociology as Social Criticism*. New York: Pantheon.
- Burgher, Thomas
1977 "Talcott Parsons, the problem of order, and analytic sociology." *American Journal of Sociology* 83:320-34.
- Chodorow, Nancy
1974 "Family structure and feminine personality." Pp. 43-66 in Michella Zimbalist Rosaldo and Louise Lamphere (eds.), *Woman, Culture and Society*. Stanford: Stanford University Press.
1978 *Mothering: Psychoanalysis and the Social*

most of his career to make significant progress on any of these questions. This failure can be linked, I believe, to his insufficient regard for the complexity of social life. In addressing the issues of institutional interrelation and the relation of institutions and personality, Habermas has simply lacked the theoretical vocabulary to distinguish the complex causal processes involved. On the issue of moral and symbolic development, moreover, he has had no substantive theory at all. Only in his most recent work has Habermas (1975) begun to surmount these difficulties. It is far from accidental that he has done so only by drawing extensively on Parsons's own theoretical system.

- Organization of Gender. Berkeley: University of California Press.
- Chomsky, Noam
1968 *Language and Mind*. New York: Harcourt Brace.
- Clark, Terry Nichols
1973 *Prophets and Patrons: The French University and the Emergence of the Social Sciences*. Cambridge, Ma.: Harvard University Press.
- Cohen, Jere, Lawrence E. Hazelrigg, and Whitney Pope
1975 "De-Parsonizing Weber: a critique of Parsons's interpretation of Weber's sociology." *American Sociological Review* 40: 229-41.
- Coser, Lewis
1956 *The Functions of Social Conflict*. New York: Free Press.
- Dahrendorf, Ralph
1958 "Out of utopia." *American Sociological Review* 64:115-27.
1959 *Class and Class Conflict in Industrial Society*. Stanford: Stanford University Press.
- Deutsch, Karl W.
1963 *The Nerves of Government*. New York: Free Press.
1964 "Integration and the social system: implications of functional analysis." Pp. 179-208 in Philip E. Jacob and James V. Toscano (eds.), *The Integration of Political Communities*. Philadelphia: Lippincott.
- Devereux, Edward C.
1961 "Parsons's sociological theory." Pp. 1-63 in Max Black (ed.), *The Social Theories of Talcott Parsons*. Ithaca: Cornell University Press.
- Dreeben, Robert
1968 *On What is Learned in School*. Reading, Ma.: Addison-Wesley.
- Durkheim, Emile
[1893] *The Division of Labor in Society*. New York: Free Press.
- Easton, David
1953 *The Political System*. New York: Knopf.
1965 *A Systems Analysis of Political Life*. New York: Wiley.
- Eisenstadt, S. N.
1964a "Institutionalization and change." *American Sociological Review* 29:235-47.
1964b "Social change, differentiation, and evolution." *American Sociological Review* 29:375-86.
1964c "Breakdowns of modernization." *Economic Development and Cultural Change* 12:345-67.
1965 "Transformations of social, political and cultural orders in modernization." *American Sociological Review* 30:359-73.
1969 *The Political System of Empires*. New York: Free Press.
1971a *From Generation to Generation*. New York: Free Press.
1971b *Political Sociology*. New York: Basic Books.
1971c *Social Differentiation and Stratification*. Glenview: Scott, Foresman.
- 1973 *Tradition, Change and Modernity*. New York: Wiley.
- Ekeh, Peter
1974 *Social Exchange Theory: The Two Traditions*. Cambridge, Ma.: Harvard University Press.
- Erikson, Erik H.
1950 *Childhood and Society*. New York: Norton.
- Evans-Pritchard, E. E.
1953 "The Nuer concept of spirit in its relation to the social order." *American Anthropologist* 55:201-14.
- Fararo, Thomas J.
1976 "On the foundations of theory of action Whitehead and Parsons." Pp. 90-122 in J. J. Loubser, Rainer Baum, Andrew Effrat, and Victor Lidz (eds.), *Explorations in General Theory in Social Science*. New York: Free Press.
- Fass, Paula S.
1977 *The Damned and the Beautiful: American Youth in the 1920s*. Oxford: Oxford University Press.
- Foss, Daniel
1963 "The world view of Talcott Parsons." Pp. 96-126 in Maurice Stein and Arthur Vidi (eds.), *Sociology on Trial*. Englewood Cliffs: Prentice-Hall.
- Friedrichs, Robert W.
1970 *A Sociology of Sociology*. New York: Free Press.
- Geertz, Clifford
1971 "After the revolution: the fate of rationalism in the new states." Pp. 357-400 in Alex Inkeles and Bernard Barber (eds.), *Stability and Change*. Boston: Little Brown.
1973 *The Interpretation of Cultures*. New York: Basic Books.
- Giddens, Anthony
1971 *Capitalism and Modern Social Theory*. Cambridge, Ma.: Cambridge University Press.
1972 "Four myths in the history of social thought." *Economy and Society* 1:357-69.
- Gintis, Herbert L.
1969 *Alienation and Power: Toward a Radical Welfare Economics*. Ph.D. dissertation, Department of Economics, Harvard University.
- Gould, Mark
1976 "Systems analysis, macrosociology, and the generalized media of action." Pp. 47-506 in J. J. Loubser, Rainer Baum, Andrew Effrat, and John Meyer Lidz (eds.), *Explorations in General Theory in Social Science*. New York: Free Press.
- Gouldner, Alvin J.
1967 "Reciprocity and autonomy in functional theory." Pp. 141-69 in N. J. Demerath and Richard A. Peterson (eds.), *Systemic Conflict and Change*. New York: Free Press.
1970 *The Coming Crisis of Western Sociology*. New York: Equinox.

- Habermas, Jürgen
 1970 *Towards a Rational Society*. Boston: Beacon.
 1973a *Knowledge and Human Interest*. Boston: Beacon.
 1973b *Theory and Practice*. Boston: Beacon.
 1975 *Legitimation Crisis*. Boston: Beacon.
- Halévy, Elie
 [1901-] *The Growth of Philosophic Radicalism*.
 1904] 3 Vols. New York: Kelley.
 1972
- Heydebrand, Wolf
 1972 "Review symposium." *Contemporary Sociology* 1:387-95.
- Hobbes, Thomas
 [1651] *Leviathan*. New York: Colliers.
 1962
- Hobsbawm, E. J.
 1973 "Karl Marx's contribution to historiography." Pp. 265-83 in Robin Blackburn (ed.), *Ideology in Social Science*. New York: Vintage.
- Hofstadter, Richard
 1952 *The Paranoid Style in American Politics and Other Essays*. New York: Random.
- Hughes, H. Stuart
 1958 *Consciousness and Society*. New York: Random House.
- Huntington, Samuel P.
 1968 *Political Order in Changing Societies*. New Haven: Yale University Press.
- Inkeles, Alex
 1971 "Continuity and change in the interaction of the personal and socio-cultural systems." Pp. 265-81 in Alex Inkeles and Bernard Barber (eds.), *Stability and Change*. Boston: Little, Brown.
- Jessop, Bob
 1972 *Social Order, Reform, and Revolution*. London: MacMillan.
- Johnson, Benton
 1976 *Functionalism in Modern Sociology: Understanding Talcott Parsons*. Morristown: General Learning Press.
- Johnson, Chalmers
 1966 *Revolutionary Change*. Boston: Little, Brown.
- Johnson, Harry M.
 1973 "The generalized symbolic media in Parsons's theory." *Sociology and Social Research* 57:208-21.
 1975 "Parsons's use of the analytic model." *Indian Journal of Social Research* 16:354-71.
- Keller, Suzanne
 1963 *Beyond the Ruling Class*. New York: Random House.
- Levine, Donald N.
 1969 "Cultural integration." *International Encyclopedia of the Social Sciences* 7:372-80.
- Levy, Marion J.
 1952 *The Structure of Society*. Princeton: Princeton University Press.
- Lichtheim, George
 1961 *Marxism: An Historical and Critical Study*. New York: Praeger.
- Lipset, Seymour Martin
 1962 "Introduction." Pp. 15-39 in Robert Michels, *Political Parties*. New York: Free Press.
- 1967 *The First New Nation*. New York: Anchor.
- 1975 "Social structure and social change." Pp. 172-209 in Peter Blau (ed.), *Social Structure*. New York: Free Press.
- Lipset, Seymour Martin and Everett Carl Ladd, Jr.
 1972 "The politics of American sociologists." *American Journal of Sociology* 78:67-104.
- Lipset, Seymour Martin and Stein Rokkan
 1967 "Cleavage structures, party systems and voter alignments." Pp. 1-64 in Seymour Martin Lipset and Stein Rokkan (eds.), *Party Systems and Voter Alignments*. New York: Free Press.
- Little, David
 1969 *Religion, Order and Law*. New York: Harper and Row.
- Lockwood, David
 1956 "Some remarks on 'the social system'." *British Journal of Sociology* 5:134-45.
- Loubser, J. J.
 1976 "General introduction." Pp. 1-22 in J. J. Loubser, Rainer Baum, Andrew Effrat, Victor Meyer Lidz (eds.), *Explorations in General Theory in Social Science*. New York: Free Press.
- Lukes, Steven
 1972 *Emile Durkheim: His Life and Work*. New York: Harper and Row.
- Marks, Stephen R.
 1974 "Durkheim's theory of anomie." *American Journal of Sociology* 80:329-63.
- Marshall, T. H.
 1965 *Class, Citizenship, and Social Development*. New York: Anchor.
- Martindale, Don
 1960 *The Nature and Types of Sociological Theory*. Cambridge, Ma.: Houghton Mifflin.
- Marx, Karl
 [1875] "Critique of the Gotha programme." Pp. 13-37 in *Selected Works of Marx and Engels*, Vol. 2. Moscow.
 [1843] "Contribution to the critique of Hegel's philosophy of right." Pp. 41-60 in T. B. Bottomore (ed.), *Early Writings*. New York: McGraw-Hill.
- Marx, Karl and Frederick Engels
 [1848] "The Communist manifesto." *Selected Works of Marx and Engels*, Vol. 1. Moscow.
 1962
- Mayhew, Leon H.
 1968a *Law and Equal Opportunity*. Cambridge, Ma.: Harvard University Press.
 1968b "Ascription in modern societies." *Sociological Inquiry* 38:105-20.
- Menzies, Ken
 1977 *Talcott Parsons and the Social Image of Man*. London: Routledge.
- Merleau-Ponty, Maurice
 [1947] *Humanism and Terror*. Boston: Beacon.
 1969
- Michels, Robert
 1962 *Political Parties*. New York: Free Press.

- Mills, C. Wright
1959 *The Sociological Imagination*. Oxford: Oxford University Press.
- Mitchell, William C.
1958 "The polity and society: a structural-functional analysis." *Midwest Journal of Political Science* 2:403-20.
1972 *The American Polity*. New York: Free Press.
- Mommsen, Wolfgang J.
1971 "Discussion of Raymond Aron's 'Max Weber and power politics.'" Pp. 109-16 in Otto Stammer (ed.), *Max Weber and Sociology Today*. New York: Harper and Row.
- Nelson, Benjamin
1949 *The Idea of Usury*. Chicago: University of Chicago Press.
- Parsons, Talcott
1928 "'Capitalism' in recent German literature: Sombart and Weber." *Journal of Political Economy* 36:641-66.
1929 "'Capitalism' in recent German literature: Sombart and Weber (concluded)." *Journal of Political Economy* 37:31-51.
1937 *The Structure of Social Action*. New York: Free Press.
1951 *The Social System*. New York: Free Press.
1954 *Essays in Sociological Theory*. New York: Free Press.
1960 *Structure and Process in Modern Society*. New York: Free Press.
1961 "Introduction to culture and the social system." Pp. 963-93 in Talcott Parsons, Kasper Naegle, Edward A. Shils, Jesse R. Pitts (eds.), *Theories of Society*. New York: Free Press.
1963 "Introduction." Pp. ix-lvii in Max Weber, *The Sociology of Religion*. Boston: Beacon.
1964 *Social Structure and Personality*. New York: Free Press.
1966 *Societies: Evolutionary and Comparative Perspectives*. New York: Prentice-Hall.
1967 *Sociological Theory and Modern Society*. New York: Free Press.
1969 *Politics and Social Structure*. New York: Free Press.
1970a "On building social system theory: a personal history." *Daedalus* 99:826-81.
1970b "Some considerations on the theory of change." Pp. 95-122 in S. N. Eisenstadt (ed.), *Readings in Social Evolution and Development*. London: Pergamon.
1971a "Belief, unbelief, and disbelief." Pp. 207-45 in Rocco Caporau and Antonio Grumelli (eds.), *The Culture of Unbelief: Studies and Proceedings*. Berkeley: University of California Press.
1971b "Comparative studies and evolutionary change." Pp. 97-139 in Ivan Vallier (ed.), *Comparative Methods in Sociology*. Berkeley: University of California Press.
1971c *The System of Modern Societies*. Englewood Cliffs: Prentice-Hall.
- Parsons, Talcott, Robert F. Bales, and Edward A. Shils
1953 *Working Papers in the Theory of Action*. New York: Free Press.
- Parsons, Talcott, and Robert F. Bales
1955 *Family, Socialization, and Interaction Process*. New York: Free Press.
- Parsons, Talcott and Gerald M. Platt
1973 *The American University*. Cambridge, Ma.: Harvard University Press.
- Parsons, Talcott and Edward A. Shils
1951 "Values, motives, and systems of action." Pp. 47-278 in Talcott Parsons and Edward Shils (eds.), *Towards a General Theory of Action*. New York: Free Press.
- Parsons, Talcott, and Neil J. Smelser
1956 *Economy and Society*. New York: Free Press.
- Parsons, Talcott, and Winston White
1969 "The mass media and the structure of American society." Pp. 241-51 in Talcott Parsons (ed.), *Politics and Social Structure*. New York: Free Press.
- Pitts, Jesse R.
1964 "Continuity and change in bourgeois France." Pp. 249-304 in Stanley Hoffman (ed.), *In Search of France*. Cambridge, Ma.: Harvard University Press.
1974 "The hippies as a contra-meritocracy." *Dissent* 21:304-16.
- Plato
1945 *The Republic*. Edited by Francis M. Cornford. Oxford: Oxford University Press.
- Pope, Whitney
1973 "Classic on classic: Parsons's interpretation of Durkheim." *American Sociological Review* 38:399-415.
- Rex, John
1961 *Key Problems in Sociological Theory*. London: Routledge.
- Rocher, Guy
1975 *Talcott Parsons and American Sociology*. New York: Barnes and Noble.
- Rueschmeyer, Dietrich
1977 "Structural differentiation, efficiency and power." *American Journal of Sociology* 83: 1-25.
- Sartre, Jean Paul
1968 *Search for A Method*. New York: Random House.
- Schwanenberg, Enno
1971 "The two problems of order in Parsons's theory: an analysis from within." *Social Forces* 49:569-81.
1976 "On the meaning of the theory of action." Pp. 35-45 in J. J. Loubser, Rainer C. Baum, Andrew Effrat and Victor Meyer Lidz (eds.), *Explorations in General Theory in Social Science*. New York: Free Press.
- Scott, John Finley
1963 "The changing foundations of the Parsonian action schema." *American Sociological Review* 28:716-35.
- Shils, Edward A.
1972 *The Intellectuals and the Powers and Other Essays*. Chicago: University of Chicago Press.

- Slater, Philip
 1961a "Parental role differentiation." *American Journal of Sociology* 67:269-308.
 1961 "Towards a dualistic theory of identification." *Merrill-Palmer Quarterly* 7:113-26.
- Smelser, Neil J.
 1959 *Social Change in the Industrial Revolution*. Chicago: University of Chicago Press.
 1962 *Theory of Collective Behavior*. New York: Free Press.
 1968 *Essays in Sociological Explanation*. Englewood Cliffs: Prentice-Hall.
 1971 "Stability, instability, and the analysis of political corruption." Pp. 7-29 in Alex Inkeles and Bernard Barber (eds.), *Stability and Change*. Boston: Little, Brown.
 1972 "Introduction." Pp. i-xli in Neil J. Smelser (ed.), *Karl Marx on Society and Social Change*. Chicago: University of Chicago Press.
 1973 "Epilogue: social structural dimensions of higher education." Pp. 389-422 in Talcott Parsons and Gerald M. Platt (eds.), *The American University*. Cambridge, Ma.: Harvard University Press.
 1974 "Growth, structural change, and conflict in California public higher education, 1950-1970." Pp. 9-141 in Neil J. Smelser and Gabriel Almond (eds.), *Public Higher Education in California*. Berkeley: University of California Press.
 1975 "Re-examining the parameters of economic activity." Paper presented at the Conference on Business and Society, Berkeley.
- Smith, Anthony
 1973 *The Concept of Social Change*. London: Routledge.
- Stark, Werner
 1962 *The Fundamental Forms of Social Thought*. London: Routledge.
- Stinchcombe, Arthur
 1968 *Constructing Social Theories*. New York: Harcourt, Brace.
- Strauss, Leo
 1953 *Natural Right and History*. Chicago: University of Chicago Press.
- Surace, Samuel
 1976 "Social theory and history in Italian modernization: unions and politics." *Studies in International Development*.
- Sztompka, Piotr
 1968 "The logic of functional analysis." *Quality and Quantity* 5:369-88.
 1974 *System and Function*. New York: Academic Press.
- Tiryakian, Edward A.
 1975 "Neither Marx nor Durkheim . . . perhaps Weber." *American Journal of Sociology* 81:1-33.
- Toby, Jackson
 1975 "The war on poverty: politics of unrealistic expectations." *Contemporary Sociology* 4:11-8.
- Trotsky, Leon
 [1938] *Their Morals and Ours*. New York: Merit.
 1969
- Turner, Jonathan H.
 1974 *The Structure of Sociological Theory*. Homewood: Dorsey.
- Turner, Jonathan H. and Leonard Beeghly
 1974 "Current folklore in the criticism of Parsonian action theory." *Sociological Inquiry* 44:47-55.
- Van Zule Slabbert, F.
 1976 "Functional methodology in the theory of action." Pp. 46-58 in J.J. Loubser, Rainer C. Baum, Andrew Effrat, Victor Meyer Lidz (eds.), *Explorations in General Theory in the Social Sciences*. New York: Free Press.
- Voegelin, Eric
 1956 *Order in History*. 2 Vols. Baton Rouge: Louisiana State University Press.
- Wallace, Walter L.
 1969 "Overview of contemporary sociological theory." Pp. 1-59 in Walter Wallace (ed.), *Sociological Theory*. Chicago: Aldine.
- Walzer, Michael
 1965 *The Revolution of the Saints*. Cambridge, Ma.: Harvard University Press.
- Weber, Max
 1952 *Ancient Judaism*. New York: Free Press.
 1954 *Law in Economy and Society*. Cambridge, Ma.: Harvard University Press.
 1958a *The City*. New York: Free Press.
 1958b *The Protestant Ethic and the Spirit of Capitalism*. New York: Scribners.
- Weinstein, Fred and Gerald M. Platt
 1969 *The Wish to be Free*. Berkeley: University of California Press.
- Whitehead, Alfred North
 1925 *Science and the Modern World*. New York: Macmillan.
- Wilson, Brian R. (ed.)
 1970 *Rationality*. New York: Harper and Row.
- Wrong, Dennis
 1961 "The over-socialized conception of man in modern sociology." *American Sociological Review* 26:183-93.

COMMUNITY STRUCTURE AND INDUSTRIAL CONFLICT: AN ANALYSIS OF STRIKE ACTIVITY IN SMSAs*

JAMES R. LINCOLN
University of Southern California

American Sociological Review 1978, Vol. 43 (April):199-220

Kerr and Siegel's (1954) theory of social conditions responsible for strike proneness is taken as a framework for understanding variations in strike activity among 78 SMSAs. Their approach suggests that industrial specialization, absentee ownership, and employment concentration are community structure variables which tend to produce an "isolated mass" of workers and consequently a high level of community-based industrial conflict. Moreover, Kerr and Siegel imply that the influence of such conditions on local strike rates is not contingent on the level and form of labor organization in the community. Contrary to this view, the present study indicates that absentee ownership and industrial specialization only produce strikes indirectly by shaping the pattern of urban unionism which, in turn, is related to variations in strike activity. On the other hand, a measure of employment concentration behaves in a fashion which is quite consistent with propositions drawn from Kerr and Siegel, although alternative explanations for the findings regarding this variable are also discussed. Two measures of urban unionism, representing the level of unionization and the degree of concentration in union organization, are used in this analysis. These have notably different effects on dimensions of metropolitan strike activity. Overall unionization increases the frequency of strikes but not their size or duration. Union concentration, on the other hand, produces just the opposite pattern of effects. The study is conceived as an investigation into how organizational properties of urban communities affect their levels of industrial conflict.

INTRODUCTION

Unlike the study of urban civil violence, a growing body of quantitative research on strikes has with few exceptions ignored the community as a locus of industrial conflict.¹ There may be a tendency among

students of the subject to assume that the city is irrelevant; that strikes are primarily reactions to national forces and localities only reflect these processes through their industrial composition. A similar conclusion regarding the absence of causal forces operating at the local level, in fact, has been reached by some students of urban racial conflict, but there it was born of a series of largely futile attempts to detect community attributes which could account for variation in disorder proneness or severity (Spilerman, 1970; 1971; 1976). No such evidence, to my knowledge, exists in the case of strikes. On the contrary, there is some evidence for Britain and France, at least, that a large fraction of the territorial variance in strikes cannot be explained solely in terms of the spatial distribution of industries

* This paper has benefited from critical readings by Michael Aiken, Jon Miller, Richard Miller, Herman Turk, and an anonymous ASR reviewer. I must admit to having followed their recommendations selectively, however. Thus, the usual disclaimers concerning their lack of responsibility for any remaining deficiencies are particularly in order. An earlier draft of this paper was presented to the Labor Studies Section of the Society for the Study of Social Problems, Chicago, 1977.

¹ Space precludes a listing of all relevant works, but a representative sampling would include: Ashenfelter and Johnson (1969); Britt and Galle (1972, 1974); Griffin (1939); Kerr and Siegel (1954); Knowles (1952); O'Brien (1965); Pencavel (1970); Rees (1954); Ross and Hartman (1960); Shorter and Tilly (1974); Skeels (1971); Snyder (1975); Snyder and Kelley (1976); Stern (1976). Of these, only the contributions by Shorter and Tilly and Stern contain systematic analyses of city data. The paper by Stern examines, as does this one, variations in strike patterns among American SMSAs. Although his concern is not chiefly with the Kerr-Siegel theory as is mine, several of the variables he considers are simi-

lar to those examined here, and his results concerning the effects of plant size and unionization are not dissimilar to those I obtain. However, this study suffers from the use of state level unionization data, governmental data pertaining only to the central city of the SMSA, and the lack of any actual measure of absentee ownership although it is regarded as a theoretically important variable.

(Knowles, 1952:207; Shorter and Tilly, 1974:255). But what facets of community structure and composition, apart from industry, affect industrial conflict is still a question to which no clear answers exist, although numerous case studies and several good speculative discussions are available (Derber et al., 1958; Goldberg, 1974; Pope, 1942; Shorter and Tilly, 1974:271; Warner and Low, 1947; Wray, 1953). The present investigation addresses the proposition that facets of urban social organization which concentrate workers locally in a homogeneous mass contribute to the level and form of industrial conflict in the urban community.

Urban Social Organization and the Kerr-Siegel Theory of Strikes

In their recent comprehensive treatment of French strikes, Shorter and Tilly (1974) argue that an analysis of the organizational structures which form the work environment, as well as those which permit the mobilization of workers against employers, is critical to an understanding of the past and present nature of strikes. This perspective is shared by the present study. Cyclical fluctuations in economic variables which influence the amount of strike activity are of less interest to me. Such patterns, in any case, are less amenable to the cross-sectional analysis to be performed here than to a cross-temporal one. My strategy might be called *ecological* in the sense that it focuses on the interplay between urban-territorial and organizational processes in structuring work relations in a fashion more or less conducive to conflict. As such, it accords with other models of community which link the mobilization of organized interest groups and the ensuing patterns of conflict or accommodation to the ecological role of the city in a national division of labor (Hawley, 1950: Chap. 11; Lincoln, 1976).

For a theory relating community organization to industrial conflict, one need only turn to a highly influential paper purporting to explain industry differences in strike propensity. As Shorter and Tilly (1974:289) point out, the logic of Kerr and Siegel's (1954) discussion of strikes:

... was in fact suited to account for territorial, not industrial differences in strike activity . . . [;] all their assertions concern the militancy-producing effects of living and working in a certain type of *place*, rather than a certain kind of industry.

Kerr and Siegel's (1954) argument was a blend of themes derived from mass society and, to a lesser extent, Marxian theory. Where employees live and work in a setting which multiplies their contacts with one another while reducing their access to other social groups and strata, their strike proneness will be high.² The argument is reminiscent of much sociological theory regarding the structural conditions under which a class or status group takes action on behalf of its interests (Dahrendorf, 1959). It is important to note, however, that the theory, broadly conceived, takes two forms: one emphasizing the collective alienation and frustration that is born of membership in an isolated mass, the other pointing to the organizational advantages which materialize when persons with similar objective interests are made easily accessible to one another, a view commonly attributed to Marx (1963:338; Bendix and Lipset, 1966). It is not hard to discern that Kerr and Siegel (1954) take the former position, and their work thus falls into that broad class of industrial conflict writing which portrays strikes as collective behavior: the spontaneous revolt by an alienated group against the larger society.³ Kerr and Siegel (1954:199) do state that union movements will be strong where workers form an isolated mass, but they do not see worker organization as a necessary condition for strike action. Both outcomes, rather, stem from a common cause:

² Aside from their discussion of the isolated mass, there are other facets to Kerr and Siegel's theory which I will not attempt to treat—particularly those pertaining to the arduousness of work and the personality types selected accordingly to different occupations. These issues receive only second emphasis in their own article and have been in large part ignored by later scholars.

³ For one such attempt to develop a "protest" model of strikes which treats them as qualitatively similar to riots, see Stearns (1974) and the critique of his approach by Amsden and Brier (1977). Also, see Shorter and Tilly's (1974:336–49) discussion of frustration-aggression as an explanatory principle in studies of strikes.

Strikes may be better explained by looking not so much at the organization as at the membership. . . . It is to the reactions of workers and not to the tactics of the leaders that we must turn for the more basic explanations.⁴

While Kerr and Siegel's paper continues to exert a considerable influence over industrial relations writing, anything approaching a rigorous test of their hypotheses regarding social isolation and industrial militancy has yet to appear. Shorter and Tilly (1974:287), however, claim to perform such a test on the basis of 1906 French data pertaining to *arrondissements* (the equivalents of U. S. counties). Arguing that Kerr and Siegel's (1954) central proposition is that industrially specialized communities should be more strike prone, they group such places into three classes: *monoindustrial*, *polyindustrial*, and *metropolitan*.⁵ Thus, the isolation of workers is judged greatest in the single industry place, less so where a mix of industries prevails. Moreover, the vast metropolitan diversity of France's three largest cities, they contend, represents: "... a clear antithesis to the isolated mass" (Shorter and Tilly, 1974:290). However, Shorter and Tilly (1974) found strike rates highest not among *monoindustrial* places but in the *metropolitan* group. On the basis of these findings, they reject the Kerr and Siegel hypothesis and argue instead in favor of a theory relating the availability of skills and resources in large cities to a high potential for worker organization.

But Shorter and Tilly's test is hardly the rigorous evaluation that such an important theory deserves. One problem in their treatment is that the hierarchy of places

they adopt is as much one of population size as of industrial heterogeneity. Kerr and Siegel, however, argued that in the largest cities some employee groups (e.g., *longshoremens*) are as socially isolated as their geographically removed counterparts in small, industrially specialized communities.⁶ But what is perhaps more important is just the inability of this crude classification of local areas to represent validly the variety of specific organizational properties of communities which are causally linked to the social concentration and isolation of workers. Besides industrial specialization, Shorter and Tilly argue that large absentee-owned plants also isolate workers from the broader community and the degree to which such units control employment will be lowest in the large metropolis. That such cities display more industrial conflict than the single-industry, "company town" seems to them to reveal the paucity of the Kerr-Siegel approach. Yet one should view skeptically Shorter and Tilly's assertions that such elements of community structure all neatly converge in the fashion they claim for their classification of French counties. No independent measures of these traits, neither their interdependence nor their contributions to local levels of industrial conflict, are presented. However, absentee ownership and the degree to which workers are concentrated in large employing establishments are important facets of urban structure and deserve serious consideration along with industrial specialization in a treatment of the local conditions responsible for strikes (see Mills and Ulmer, 1970). Let us examine each in greater detail.

When major employers are national corporations with headquarters outside the local community, Kerr and Siegel (1954:192-3) write:

... the workers are as detached from the employer as from the community at large. . . .

⁴ If Kerr and Siegel attach low importance to union leadership processes in their discussion of the conditions producing strikes, other writers, dealing specifically with industrial relations in the United States, have reached different conclusions. Ashenfelter and Johnson (1969), for example, see the interaction between union leaders and rank and file as having strong implications for the pattern of strike activity. See also Britt and Galle (1972:54).

⁵ Prior to the appearance of Kerr and Siegel's paper, Knowles (1952:209) observed that the degree of industrial specialization or diversity was a community attribute which might affect the social integration of workers in an area and thus their tendency to strike.

⁶ Shorter and Tilly (1974:289) acknowledge this social isolation dimension to Kerr and Siegel's theory but, for reasons that are not entirely clear, treat it as unworthy of serious consideration. One should note that despite the wide diffusion of the proposition that large cities produce isolation and alienation, there is relatively little good evidence to support it (see Fischer, 1973).

The strike for this isolated mass is a kind of colonial revolt against far-removed authority.

The notion that absentee ownership shapes the form of urban political conflict is a familiar theme in the literature on community power and decision making (Aiken, 1970; Pellegrin and Coates, 1956; Schulze, 1958). But its implications for local patterns of industrial conflict also have received wide attention. Warner and Low's (1947) account of the shoe factory strike in Yankee City identifies absentee ownership as the key factor determining the course taken by that particular struggle. The antipathy of community residents toward the factory's managers and their sympathy for the union were aroused when control over a once locally owned industry passed to outside capital. In a case study of a Southern textile town, Pope (1942) similarly discusses the role of absentee ownership in eliciting support from the local citizenry for labor organizing efforts. In contrast, the Lynds' (1937) postdepression description of Middletown documents the control which a powerful local family of owners exercised not only over the social and economic institutions of the community, but over its perceptions of the legitimacy of unionism, so that a fledgling union found itself confronted with a solid wall of community opposition. Note, however, that these accounts do more than merely buttress Kerr and Siegel's claim that absentee ownership stimulates strikes because it contributes to worker isolation. Their main stress is rather on its role in shaping community opinion in a direction favorable to labor's side in industrial disputes. The influence of such opinion in determining the degree of labor militancy is a well-established theme in the industrial relations literature (Ashenfelter and Johnson, 1969; Ashenfelter and Pencavel, 1969:429; Goldberg, 1974).

Kerr and Siegel mention only in passing the size of employing organizations as a variable associated with strike activity. Yet the scale of economic enterprise is at least as relevant to the question of whether workers form an isolated mass as is industrial specialization: a large plant

concentrates workers in a confined site and promotes their cooperation via a complex division of labor, while it substitutes formal bureaucratic controls for personal supervision by employers (Blauner, 1964:22-3). Moreover, the size distribution of employing units in a particular community may affect industrial relations in ways which supercede the processes internal to individual organizations. A multitude of small, independent firms in the community, associating employers with relatively small groups of workers, produces a community-wide pattern of intensified cross-class interaction. At the same time, persons who occupy similar positions in the local system of economic relations tend to be separated by their assignment to different organizational units. But in the prototypical company town, organizational cleavages are minimized and class becomes the chief dimension of social structure (see Stinchcombe, 1965).

Many writers, of course, have pointed to the size of the work organization as a factor contributing to the alienation of employees and thus their propensity to strike (Britt and Galle, 1974; Cass, 1957; Eisele, 1974; Ingham, 1970; Revans, 1956). Shorter and Tilly (1974:227-35), too, note the existence of a correlation between plant size and strike activity, but their explanation for this tendency is most curious in light of their previous denunciation of Kerr and Siegel's ideas, for the argument they advance at this point is hard to distinguish from the thrust of the isolated mass theory. The high incidence of strikes in large enterprises according to Shorter and Tilly is not, as Marx and Engels implied, the consequence of better communication and more frequent interaction among employees. On the contrary, they contend that the small work site, because it allows for intense, informal interaction within a closely knit group of workers is the most fertile ground for developing labor organizations. The strikes that erupt from large establishments, they argue, are spontaneous reactions to the alienating and dehumanizing conditions which characterize that kind of work environment.

The present investigation is not seen strictly as a test of Kerr and Siegel's

theory of structural conditions producing strikes because it deals solely with metropolitan areas as units of analysis and thereby fails to examine those work settings they saw giving rise to the isolated mass, which appear in rural regions or in pockets of the largest urban centers. But it adopts from Kerr and Siegel, Marx, Lipset (1963), and others the proposition that social and economic systems which concentrate workers in large, undifferentiated groups will be subject to high levels of industrial and political conflict. Furthermore, a major concern here is to determine whether the relation between concentration and conflict is mediated by the level and form of unionization in the community—a possibility suggested by Marx's stress on the intervening role of labor organization. This model positing indirect effects of this sort will, in the course of the analysis, be contrasted with a strict interpretation of the Kerr and Siegel theory which treats strikes as more or less immediate reactions to social isolation with unions doing little to articulate this relationship. Of course, the effects of unionism on the pattern of strike activity in urban communities is itself of interest and will be the topic of much of the following discussion. At its most general level, finally, this study evaluates the proposition that community structure affects local industrial conflict in ways not reducible to community variations in the presence of differentially strike-prone industries.

MEASUREMENT AND DESIGN

In 1965, the U.S. Bureau of Labor Statistics (1966) drew a probability sample of 84 Standard Metropolitan Statistical Areas from the 221 SMSAs then in existence. Of the approximately 52,000 establishments which fell within this area sample, a sample of 13,000 establishments was taken. The survey of these employing units generated estimates of the prevalence of union contracts for the 84 SMSAs. This measure of local unionization is used in the present study; thus, the BLS sample delineates the set of SMSAs to be considered here. However, data on several important variables were not available for six small SMSAs in the origi-

nal sample so the analysis is confined to the remaining 78. These areas have been measured on the following 12 variables.

1. *Industrial Specialization* (INDSP). As Shorter and Tilly's treatment suggests, the degree to which an area is industrially specialized is clearly a factor in Kerr and Siegel's conception of features of the work environment which produce an isolated mass of workers. A commonly used index of industrial diversification is that devised by Gibbs and Martin (1962; also see Gibbs and Poston, 1975; and Lieberman, 1969). Its purpose is to describe the extent of employment dispersion among a fixed set of industrial categories for a given place. Although the Gibbs-Martin measure is usually presented as an index of diversification so that high values mean that multiple industries receive equal shares of employment, present theoretical constructions are better approximated by reversing its direction, therefore giving high scores to metropolitan areas with highly specialized economies. The measure is computed on a 26-item industrial classification of employment.⁷

2. *Absentee Ownership* (ABSOWN). If one defines the extent of absentee ownership as the amount of employment in establishments whose owners reside in another locality, the present measure is subject to error in two respects. First, it pertains only to manufacturing industries, and, second, only to large corporations. It is specifically the estimated proportion of SMSA employment in plants owned by firms listed in the 1963-1964 *Fortune Plant and Product Directory* (Fortune, 1963) whose headquarters lie outside the SMSA in question. Since the majority of strike activity is in manufacturing, and these firms accounted for approximately 70% of the nation's output of manufactured goods in 1963, the measure is perhaps less limited than it might initially appear. Moreover, a strong case can be made that it is precisely this kind of control over local employment—by the largest, most centralized corporations—

⁷ The industry data originated with the Office of Business Economics, U. S. Department of Commerce. The specific industries in the 26-item list are presented in the appendix.

which contributes most to local employees' sense of isolation from employers.

3. *Employment Concentration* (EMP-CON). The U. S. Bureau of the Census (1967) File, *County Business Patterns*, supplies data on the employment-size distribution of establishments in all industries but government and railroad transportation for counties and SMSAs in the United States. To summarize this distribution in a fashion assigning high values to metropolitan areas with highly concentrated employment, a gini coefficient was calculated over it for each SMSA. The data pertain to 1967.⁸

4. *Unionization* (UNIZ). This measure, drawn from the BLS survey, represents the extent of unionization in the metropolitan labor force. It is the estimated proportion of workers in establishments which had a labor-management agreement in effect covering the majority of employees.⁹

5. *Union Concentration* (UCON). Aside from the degree of labor force unionization, numerous writers have identified the extent to which worker organizations are centralized and bureaucratized as a pattern having distinct implications for strike activity (Britt and Galle, 1972; Ingham, 1974; Ross and Hartman, 1960; Shorter and Tilly, 1974:165-73). Seidman (1965), furthermore, has paid special attention to the degree of centralization in *metropolitan* union organization as a circumstance shaping local patterns of conflict. Implied in his and other treatments is the notion that centralization in unionism, by increasing the capacity to mobilize workers and resources, has the potential for increasing the scale and duration of conflict events. The sheer number of work stoppages is seen less frequently to reflect the organizational capacities of

workers (Knowles, 1952:146; Skeels, 1971). In this paper, concentration of metropolitan union power is measured in much the same way as EMPCON and derives from the same data base. It, too, is a gini coefficient, tapping concentration in the distribution of employees over organizational units where the latter in this case are labor associations. Thus, it should attain high values in SMSAs where most paid union staff are attached to a relatively few large labor organizations.¹⁰

6. *Number of Work Stoppages* (STRIKES).

7. *Number of Workers Involved* (STRIKERS).

8. *Number of Man-Days Idle* (MANDAYS). The Bureau of Labor Statistics (1964-1971) publishes these three sets of figures regularly pertaining to yearly strike activity in SMSAs. The data are totals computed over all industries; no industry-specific figures for urban areas are available. I have aggregated the annual SMSA totals over the period, 1963-1969. The number of strikes in the United States was the lowest since World War II in 1963. That year, however, began a period of uninterrupted increase which ended in 1971, when strike activity dropped sharply in response to President Nixon's program of wage and price controls (U. S. Bureau of Labor Statistics, 1973). By summing the yearly strike totals over this interval, one averages out fluctuations due to short-run business and contract expiration-renewal cycles. The latter, in particular, can produce sharp shifts from one year to the next in certain highly specialized urban economies. The number of Detroit workers involved in strikes in 1964, for example, was seven times greater than the previous year (from 15,500 to 114,000) because of strikes induced by contracts ending in 1964 with the four leading auto makers (U.S. Bureau of Labor Statistics, 1965). The distributions on the strike variables are highly skewed in this sample,

⁸ The employment-size classification used by *County Business Patterns* is 1-3, 4-7, 8-19, 20-49, 50-99, 100-249, 250-499, 500+.

⁹ The BLS survey distinguished plant from office workers and presented separate unionization data for these groups. The present data pertain only to plant workers. The industries represented are: manufacturing, transportation, communication and utilities, finance, insurance and real estate, wholesale and retail trade and services. The same measure of unionization has been used previously in studies by Aiken and Alford (1970) and Hill (1974).

¹⁰ These data are published in *County Business Patterns* (U.S. Bureau of the Census, 1967) under the SIC code 863: nonprofit membership organizations for the advancement of labor interests. Included are employees' associations, labor unions, and other labor organizations.

and a log transformation improves the ability of a linear combination of independent variables to predict them. The results to follow are based on the logged variables.

The measures described above are suggested by the key theoretical question posed by this investigation: how do facets of urban industrial organization which presumably concentrate workers in homogeneous groups affect strike activity in metropolitan communities? The remaining variables are, in general, less central to that issue; but their consideration is warranted by the prospect that were they omitted, parameter estimates for the theoretical variables might prove to be biased.¹¹

9. *Labor Force Size (SIZE)*. There are, of course, certain theoretical issues having to do with the isolating effects of community size, but a clear prediction regarding the role of size is not easily derived from Kerr and Siegel. It is even less easy to separate the necessary component of the correlation between SMSA size and strike activity from the causal one. That is, the quantity of conflict tends to rise with the size of the labor force simply because of rising opportunities at the industry, plant, and employee levels, and this tendency must be held constant to ascertain the effects of theoretically important but size dependent variables.¹² At the same time, the strike proneness of a fixed number of workers, industries, or establishments may change with the size of the SMSA for theoretical reasons of the kind discussed previously. I will attempt, in passing, to give attention to both kinds

of size effects, but the bulk of the analysis will treat size merely as a control. The measure, then, is the total employed of the SMSA, logged to normalize its distribution. Total employed is the relevant population for present purposes so it is preferred to total population. This decision, however, is of little consequence, for the correlation between them is .99.

10. *Proportion Employed in Manufacturing (MANU)*. Because observed associations between organizational variables and strikes may spuriously reflect SMSA industry composition, some method of controlling for industry variation is mandatory. I adopt two strategies here. First, SMSA specialization in manufacturing is included in the regression equations to be estimated.¹³ The configuration of industrial specialization, absentee ownership, and employment concentration is central to the Kerr-Siegel theory. It also describes the manufacturing city, however, where one would expect to find high levels of unionization and strike activity simply because of the collective bargaining patterns which distinguish this industry nationally. Including percent employed in manufacturing as an additional regressor should eliminate bias of this sort. Finally, the restriction of the absentee-ownership data to manufacturing would be reason alone to hold constant this industry's share of metropolitan employment.

The second method I use to adjust for industry composition is to standardize both the strike measures and EMPCON. The combination of these various procedures should make one reasonably confident that SMSA industry differences are not responsible for the results obtained.

11. *Proportion Nonwhite (NWHITE)*. Seidman (1965) has advanced a view similar to that of Kerr and Siegel concerning the relative absence of social divisions among employees in a community as a circumstance affecting local levels of labor militancy. But his attention, unlike theirs, is drawn to the racial and ethnic diversity of the area as a factor impairing worker solidarity and thus depressing the ability to mount strikes (see Kornblum,

¹¹ Unless otherwise indicated, the data regarding these variables were obtained from the 1960 U. S. Census of Population.

¹² One can view the occurrence of strikes in SMSAs as samples from a binomial distribution where the number of trials (N) is proportionate to the size of the metropolitan labor force. The mean of the binomial, $E(X)$, equals Np where p is the probability of a conflict event. It is thus easy to see how p could remain constant over SMSAs while the amount of strike activity varied markedly, simply because of variation in the size of the population at risk. A systematic attempt to establish a baseline stochastic model for the expected number of racial disorders under the assumption that the only relevant variable is the size of the black population can be found in Spilerman (1970; 1971).

¹³ The manufacturing data are from the Office of Business Economics source (see fn. 7).

1974). Like the Kerr-Siegel theory, this argument is reminiscent of certain longstanding themes in political sociology (e.g., Alford, 1963). But it implies a different approach to these issues from that of Kerr and Siegel which addresses the organization of communities and industries rather than the ethnic diversity of the work force. One can hardly ignore the question of urban ethnic composition, however, especially since the metropolitan distribution of the nonwhite labor force is systematic in regard to industry, SMSA size, plant size, and other variables central to this inquiry (see Thompson, 1965). Therefore, the proportion nonwhite will be included in the regressions to follow.

12. *Unemployment (UNEMP)*. Economists have paid special attention to the level of labor market tightness as a business cycle correlate affecting the strength of unions and the probability of strikes. Rees (1954:218) puts the matter succinctly:

The principal economic factor affecting union behavior is the state of the labor market—the amount of employment available. Rising employment and improving business conditions offer the unions a variety of strategic advantages. [The employer's] ability to replace strikers with nonstrikers diminishes as employment rises, and the strikers have an increased chance of obtaining employment elsewhere if the employer succeeds in replacing them.

Studies of the effects of unemployment and other economic conditions on union growth and industrial conflict have been conducted almost uniformly with time-series data (Ashenfelter and Johnson, 1969; Ashenfelter and Pencavel, 1969; Weintraub, 1966). Cross-sectional research on variations among urban labor markets in their levels of chronic unemployment has not extended to topics of industrial conflict (an exception is Stern, 1976). In the present analysis, the SMSA unemployment level is measured as the mean of the January rates reported by the U. S. Department of Labor over the period 1963–1967.¹⁴

The level of wages is another variable which is routinely considered in longitudinal research on strike activity. I will not formally present regression equations containing this variable as a regressor because of the obvious problem of establishing causal priority. However, to determine whether the parameter estimates for the measures discussed above are seriously biased by the omission of a wage indicator, the regressions of the strike variables in Tables 2 and 4 were expanded to include 1960 mean wage income. The coefficients calculated for this measure proved not to be significantly different from zero—those in standard form, in fact, were miniscule—nor were the effects of other variables in the model disturbed. It seems safe to conclude, then, that the results to be presented shortly are not distorted by the absence of a measure of the metropolitan wage level.

ANALYSIS AND RESULTS

The theory under consideration which links metropolitan strike activity to facets of metropolitan labor and industrial organization is portrayed as the diagram in Figure 1. A block of exogenous variables indicative of the industrial organization, labor force composition, and economy of the SMSA determines the levels of two union organization variables, unionization and union concentration. The correlated residuals relating these measures reflect the lack of a theory stipulating the form and direction of any causal connection between them. Both these sets of variables, in turn, affect the levels of the three strike dimensions—correlated residuals again implying possible causal but (for the moment) unanalyzed relationships. Should residual correlations of this sort prove small, however, the appropriate conclusion would be that such effects are relatively trivial as compared with those of causally antecedent conditions (Duncan, 1970).

The zero-order correlations, means, and standard deviations for all variables are

force and have not been seasonally adjusted. They are made available in a series of publications by the Bureau of Employment Security, U. S. Department of Labor (1963–1967).

¹⁴ These unemployment rates are numbers of unemployed as a percentage of the metropolitan work

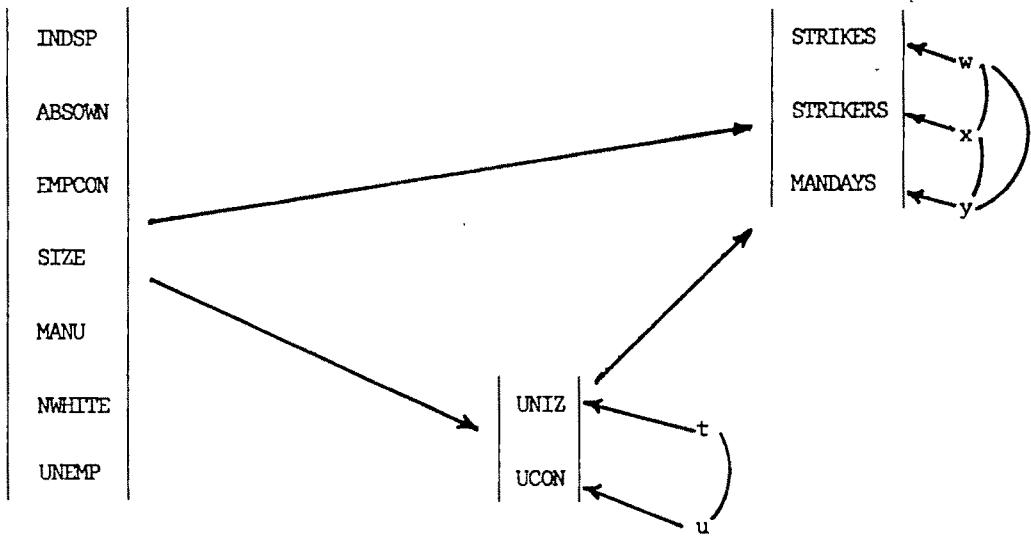


Figure 1. Path Diagram of Causal Relationships among Urban Structure, Union Organization, and Strike Variables

presented in Table 1. Not surprisingly, high positive correlations appear among the three strike indicators. The correlation between the union measures, UNIZ and UCON, is also positive but considerably weaker at .388. It further is worth noting the positive correlations among manufacturing specialization, the industrial organization measures, and the union organization measures. This pattern is consistent with the expectation that the complex seen here as generating strikes also describes the manufacturing community. The anticipated high correlations between the strike measures and SIZE are also in evidence.

Table 2 presents the ordinary least-squares estimates for the fully recursive model depicted in Figure 1. Attention should be focused immediately on some fairly clear patterns among these results which speak directly to the central theoretical question of the study. Recall that Kerr and Siegel's discussion of facets of local industrial organization giving rise to worker isolation took explicit account of industrial specialization and absentee ownership. While employment concentration might be viewed similarly, their treatment appears to give it less emphasis. But the mechanisms whereby INDSP and ABSOWN seem to be linked to strike activity are not those suggested by the Kerr-Siegel theory, although they would

seem to be those anticipated under the alternative "indirect effects" model. The direct effects of absentee ownership and industrial specialization are nonsignificant, trivial, and even have the wrong signs. But because ABSOWN and INDSP are positively related to different union organization variables (UNIZ and UCON, respectively), while increments in the latter are directly paired, in turn, with increases in the amount of strike activity, the proposition that industrial specialization and absentee ownership are associated with higher urban strike rates is supported. But contrary to Kerr and Siegel, it is clear that these features of community industrial organization dispose the area toward industrial conflict only to the degree they produce a broader and administratively more centralized labor movement.

The picture is quite different, however, in the case of employment concentration. The effects of this industrial organization variable are entirely direct. The coefficients relating EMPCON to all three strike indicators are positive and significant. At the same time, its effects on UNIZ and UCON are not significant, nor, in the case of UNIZ, is the direction of the relation the expected one.

Turning to the control variables, one finds that MANU has a strong influence on UNIZ but does not affect the other

Table 1. Zero-Order Correlations, Means and Standard Deviations (N=78 SMSAs)*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	\bar{X}	SD
1 INDSP																	.093	.018
2 ABSOWN	087																.111	.069
3 EMPCON	447	058															.752	.020
4 UNIZ	379	241															.649	.194
5 UCON	388	526	561														.509	.088
6 STRIKES	461	609	681	577													5.009	1.345
7 STRIKERS	953	847	593	261	001												6.182	1.580
8 MANDAYS	940	694	326	044	116	711											8.832	1.729
9 SIZE	743	339	048	201	768	787	792	266									12.332	1.032
10 MANU	139	088	116	978	972	962	188										.255	.111
11 NWHITE	-208	186	242	270	301	492											.108	.100
12 UNEMP	-322	033	031	011	629												.050	.014
13 EXPSTK (strikes)	151	139	154	294													4.840	1.012
14 EXPSTK (strikers)	994	993	251														6.289	1.062
15 EXPSTK (mandays)	998	262															9.037	1.074
16 SOUTH	-284																.269	.446

* Decimal points of correlation coefficients have been omitted.

Table 2. Ordinary Least-Squares Regression Estimates of Equations for Union Organization and Strike Measures (N=78 SMSAs)

Dependent Variables (Intercepts)	Predetermined Variables										R ²	R ² -
	Unstandardized Regression Coefficients (Standard Errors)											
	INDSP	ABSOWN	EMPCON	SIZE	MANU	UNEMP	NWHITE	UNIZ	UCON			
UNIZ (0.981)	-0.607 (1.256)	0.759 (0.328)	-2.091 (1.331)	0.068 (0.021)	0.923 (0.251)	2.945 (1.272)	-0.056 (0.187)				.469	.424
UCON (-0.578)	1.362 (0.561)	-0.225 (0.146)	0.290 (0.594)	0.059 (0.009)	0.066 (0.112)	-0.517 (0.568)	0.016 (0.084)				.492	.449
STRIKES (-20.752)	-3.295 (8.494)	-2.161 (2.229)	26.373 (8.792)	0.388 (0.175)	-2.733 (1.774)	-3.231 (8.618)	0.128 (1.213)	2.906 (0.777)	1.267 (1.741)		.552	.500
STRIKERS (-24.922)	-5.345 (7.953)	-1.943 (2.087)	29.893 (8.233)	0.526 (0.164)	-2.448 (1.661)	-10.431 (0.164)	0.591 (1.135)	3.175 (0.728)	3.652 (1.630)		.715	.682
MANDAYS (-27.830)	-10.152 (6.975)	-1.837 (1.830)	35.334 (7.220)	0.559 (0.144)	-2.530 (1.457)	2.299 (7.077)	1.050 (0.996)	2.638 (0.638)	5.970 (1.430)		.817	.796
Standardized Regression Coefficients ^b												
Residual r's												
UNIZ	-.055	.269*	-.214	.362***	.531***	.219*	-.029				r ₁₀ = .091	
UCON	.272*	.175	.065	.687***	.083	-.084	.019				r ₁₀ = .934	
STRIKES	-.043	-.110	.389**	.298*	-.226	-.095	.010	.418***	.083		r ₁₀ = .680	
STRIKERS	-.060	-.084	.375***	.344**	-.172	-.095	.037	.389***	.204*		r ₁₀ = .807	
MANDAYS	-.104	-.073	.405***	.334***	-.163	.019	.061	.295***	.305***			

^a R² is R² adjusted for degrees of freedom.^b One, two, and three asterisks indicate significance levels of .05, .01, and .001, respectively.

endogenous variables. NWHITE influences nothing. The SMSA unemployment rate yields a coefficient twice its standard error only in the case of UNIZ, but its sign is at odds with the argument from Rees that unemployment depresses union strength. Some (Olson, 1971:39) contend that while antecedent upturns in unemployment may depress recruitment to unions, strong unions may, in turn, aggravate unemployment by restricting access to jobs. Thus, the positive slope found here might indicate a misspecification of a fundamentally reciprocal causal process. Also possible, however, is that unemployment acts in this model as a proxy for a set of unmeasured attributes of the specialized manufacturing economy which, while highly unionized, is also subject to high cyclical unemployment.

The effects of SIZE are positive and significant on all five endogenous variables. That large urban areas are most likely to have a strong and centralized labor movement is consistent with Shorter and Tilly's (1974:275) view that they offer a range of resources and militancy-encouraging ideas to unionists. Moreover, the cultural context of the metropolis is apt to produce more tolerance for labor organizing than is found in smaller communities.¹⁵

Setting aside those components of the correlation between SMSA size and strike proneness which might be attributed to increasing opportunities and the mediating role of union organization, it is doubtful that the size of the urban community has any significant causal influence on industrial conflict. One approach to this problem is to form the ratio of each strike measure to SIZE and compute the correlation between this ratio and SIZE (see Fuguitt and Lieberman, 1974). The resulting correlations are: .292 for STRIKES, .418 for STRIKERS, and .531 for MANDAYS. These outcomes would seem to be in accord with those presented by Shorter and Tilly (1974:277) for France: strike activity as a ratio to labor force size is higher

in larger cities. However, this inference is demolished when the other predetermined variables in the equations for strike activity are controlled, for the partial correlations are: STRIKES (.017); STRIKERS (.030); and MANDAYS (-.097).

The partial correlation, r_{iu} , relating the residuals in UNIZ and UCON when the variance explained by their mutual antecedents is removed, is not significant at the 5% level. However, the residual correlations among the strike indicators remain quite high and point to a conclusion one might have drawn at the outset. Specifically, besides the dependence of different strike dimensions on a set of common causal conditions, causal ties exist among the strike variables themselves. Indeed, this pattern of partial correlations is consistent with a specification which places STRIKES antecedent to STRIKERS and both antecedent to MANDAYS. As an extension of the previous model, the path diagram in Figure 2 is thus presented with causal parameters estimated for the relations among the strike variables as well as for their joint dependence on the urban structure and union variables. In this treatment only variables with significant direct effects on strike activity are considered, and of their paths only those with coefficients twice their standard errors are presented. The chief value of this demonstration is that it underscores a pattern hinted at in Table 2: that UNIZ affects the incidence of strikes, while UCON determines their size and duration. These inferences are warranted since the residual variation in STRIKERS with STRIKES held constant must be due to size, and the variation in MANDAYS net of the other strike dimensions must be due to duration. These findings support the proposition that concentration in labor organization contributes to a capacity to stage large and long stoppages but has little impact on their frequency per se. Note, however, that significant effects of EMPCON and SIZE are still present for each dimension of conflict.

Nonlinear Models

Another possibility to consider here is that the functional form of the relations

¹⁵ However, too much importance should not be attached to the strong effect of SIZE on UCON, for UCON is somewhat constrained to low values in small SMSAs where unions are small and few in number.

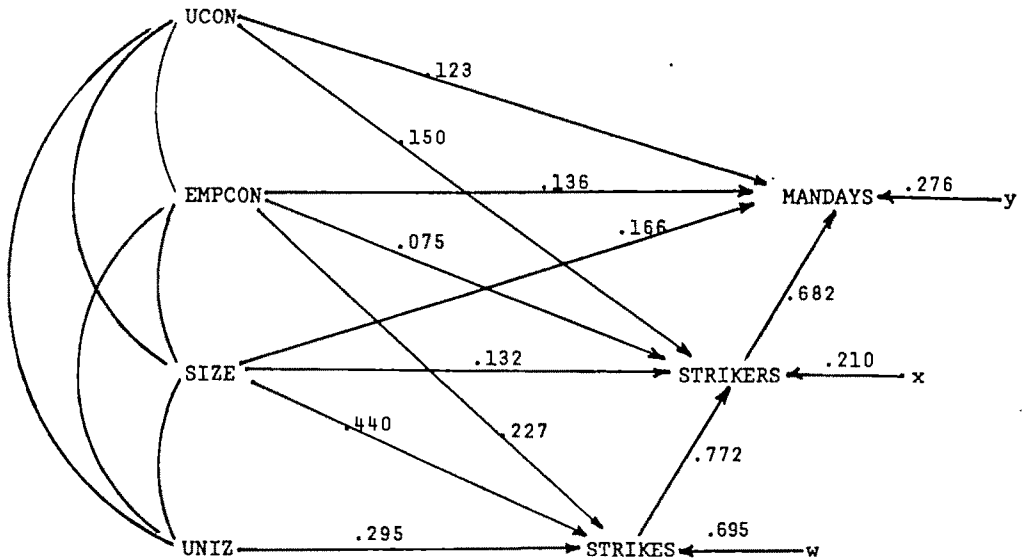


Figure 2. Path Model of Variables Affecting Strike Activity in SMSAs

tying strikes and union organization to their community-level causes is a non-linear one. One interpretation of Kerr and Siegel's discussion with regard to community size, for example, is that the effect of size is curvilinear: strike proneness should be high in the small area which concentrates workers in a limited space but also in the large city which promotes segregation and isolation of social classes. By dividing the distribution on total employment into quartiles and treating this variable as a categorical measure, a test for nonlinearity which evaluates the ability of this transformation to explain greater variance than that achieved under the quantitative form of the variable was performed. However, no increment to the R^2 was obtained by this method and the hypothesis of curvilinearity was rejected.

Certain arguments regarding the ways union organization mediates the impact of urban structure variables on strike activity suggest other departures from the linear model, especially interaction effects. Knowles's (1952:210) contention that "unrest takes the form of strikes only if workers have some degree of cohesion," points to one such interpretation. Namely, alienating features of the work setting will produce strikes as behavioral outcomes as opposed to such individualistic responses as absenteeism and high turnover rates only if workers are well organized. Inter-

preting organization as unionization, Knowles's statement suggests that the effects of community structure on strike activity should become more pronounced as the urban work force becomes more unionized. But the thrust of Kerr and Siegel's argument points to a quite different pattern. Where labor-management relations are handled through powerful unions, the isolated mass is unlikely to produce strikes. Grievances which form the basis for strikes engineered by the officialdom of well-established labor organizations tend to revolve around the more easily defined and negotiated issues of wages, pensions, and job security (see Snyder, 1975). The kind of rank-and-file revolt which Kerr and Siegel appear to have in mind should be most likely where unions are weak and fragmented.

In fact, this prediction from the Kerr-Siegel theory receives a measure of support in the present data. Two sets of interaction terms were formed, being the products of UNIZ and UCON, respectively, with the exogenous variables. The clearest pattern of interaction pertains to EMPCON. As Table 3 shows, of six coefficients estimated for product terms formed by EMPCON and the union measures, four of these produce highly significant F-ratios and have negative signs. These results are compatible with an argument which states that metropolitan

Table 3. F-Ratios for Interactions between UNIZ and UCON and Other Predetermined Variables *

Dependent Variables	Interaction Terms						
	UNIZx INDSP	UNIZx ABSOWN	UNIZx EMPCON	UNIZx SIZE	UNIZx MANU	UNIZx UNEMP	UNIZx NWHITE
STRIKES	0.003 (-)	0.051 (-)	0.000 (+)	0.461 (-)	0.140 (-)	0.776 (-)	0.075 (-)
STRIKERS	0.061 (+)	0.038 (+)	0.040 (-)	0.312 (-)	0.001 (-)	2.157 (-)	0.138 (-)
MANDAYS	0.028 (-)	0.011 (-)	5.369* (-)	0.960 (-)	2.907 (+)	3.989* (-)	0.206 (-)
Dependent Variables	UCONx INDSP	UCONx ABSOWN	UCONx EMPCON	UCONx SIZE	UCONx MANU	UCONx UNEMP	UCONx NWHITE
	INDSP	ABSOWN	EMPCON	SIZE	MANU	UNEMP	NWHITE
STRIKES	0.773 (+)	1.512 (+)	6.167* (-)	3.838 (+)	0.002 (+)	1.928 (+)	1.168 (+)
STRIKERS	1.218 (+)	2.544 (+)	5.969* (-)	3.179 (+)	0.038 (-)	1.866 (+)	0.418 (+)
MANDAYS	0.264 (+)	1.646 (+)	11.556** (-)	1.164 (+)	0.125 (+)	0.653 (+)	0.601 (-)

* F-ratios produced by estimating two equations for each strike measure: (1) one including all regressors in Table 2 plus the seven products given in the first half of the present table above; (2) the same base equation plus the seven products in the second half of the present table. Signs in parentheses indicate the directions of the estimated coefficients on the interaction terms. One, two, and three asterisks indicate significance levels of .05, .01, and .001, respectively.

areas with strong, centralized unions are less likely to experience strikes caused by high employment concentration.

Adjusting for Industry Composition

Although MANU proved to have little direct influence on metropolitan strike activity, one might contend that such a heterogeneous industry hardly constitutes an adequate measure of SMSA industrial composition. Therefore, the hypothesis that industrial variation underlies the results obtained thus far cannot yet be ruled out. A more rigorous approach is to standardize strike rates for industry composition. Since industry-specific strike rates are not published for SMSAs, the only option is indirect standardization (see Bogue, 1969:24). By applying the schedule of industry-specific rates for the nation to the industry composition of each SMSA (defined as the 26-item classification on which INDSP was constructed), expected levels of the strike variables were generated for each of the 78 SMSAs.¹⁶ Since

these expected measures (EXPSTK) depend not only on industry strike rates but on SMSA employment size, each was simply substituted for SIZE in the equation predicting its observed counterpart. MANU is also excluded from these equations since its function in the analysis is duplicated by EXPSTK. As Table 4 shows, however, the results are practically identical to those obtained previously. The differences between the two sets of R²'s are slight, and only minor changes are registered in the coefficients, although the effects of EMPCON prove to

(1965-1971) which provided the SMSA figures are also the source for the industry strike data at the national level. These data were likewise summed over the same seven year period, and the BLS industry classification was condensed to yield the 26-item scheme mentioned earlier. Strike rates by industry were then derived by calculating the ratio of strike activity in the industry to its total employed. For each SMSA and strike indicator, then, the formula for EXPSTK, the expected quantity of strike activity, is $\sum_{i=1}^{26} a_i X_i$ where a_i is the national strike rate

for the "i"th industry, and X_i is the amount of SMSA employment in that same industry.

¹⁶ The same Bureau of Labor Statistics reports

Table 4. Regressions of Observed Strike Measures on Industry-Based Expected Strike Measures and Other Predetermined Variables (N=78 SMSAs)

Dependent Variables (Intercepts)	Predetermined Variables								R ^a	R ^b
	Unstandardized Regression Coefficients (Standard Errors)									
	INDSP	ABSOWN	EMPCON	EXPSTK ^a	UNEMP	NWHITE	UNIZ	UCON		
STRIKES (-10.873)	-12.215 (7.055)	-2.880 (2.126)	17.719 (7.553)	0.328 (0.175)	-1.563 (8.712)	0.537 (1.216)	2.526 (0.720)	1.578 (1.738)	.529	.482
STRIKERS (-14.888)	-13.437 (6.716)	-2.951 (1.995)	21.707 (7.082)	0.452 (0.162)	-8.168 (8.182)	1.000 (1.144)	2.747 (0.689)	3.910 (1.660)	.698	.668
MANDAYS (-17.834)	-18.683 (5.892)	-2.926 (1.753)	25.686 (6.362)	0.490 (0.142)	4.628 (7.232)	1.537 (1.011)	2.160 (0.610)	6.255 (1.455)	.803	.784
Standardized Regression Coefficients ^c										
STRIKES	-.160	-.147	.261*	.247	-.017	.040	.364***	.104		
STRIKERS	-.150*	-.128	.273**	.304**	-.074	.063	.337***	.219*		
MANDAYS	-.191**	-.116	.295***	.305***	.038	.089	.242***	.319***		

^a The mnemonic, EXPSTK, refers to a different variable in each equation, since it is the expected level of the dependent variable.

^b R² is R² adjusted for degrees of freedom.

^c One, two, and three asterisks indicate significance levels of .05, .01, and .001, respectively.

be somewhat smaller in this new set of equations, and the negative influence of INDSP a bit more pronounced.

Since EMPCON is the only feature of urban industrial organization having a significant direct effect on conflict levels, a proper interpretation of its role seems especially important in evaluating the kinds of propositions advanced by Kerr and Siegel. As a final attempt to rule out an explanation based on industry composition, then, EMPCON itself was standardized for industry. *County Business Patterns* provides a detailed industry breakdown of the employment size distribution of organizational units in localities. These industries were aggregated to approximate the 1960 Census condensed classification (U.S. Bureau of the Census, 1960), producing 37 categories. The CBP size-by-industry matrix was then aggregated over the 224 SMSAs as defined in 1967 to produce a set of values for the nation which would serve as a standard population. Following the procedures outlined by Kitagawa (1955), it was then possible to generate three establishment size distributions for each SMSA, each of which was subsequently summarized by a gini coefficient.¹⁷ Thus, EMPCON was partitioned into three components: (1) a directly standardized or within-industry component; (2) a between-industry component, capturing only that portion of

metropolitan variation in the size distribution of units which exists between industries; and (3) a component attributable to the covariance of (1) and (2). The last term is not relevant for present purposes, but it is useful to compare the effects of the first two ginis on strike activity. Thus, for each strike variable, two equations were estimated: one substituting for the original total EMPCON its within-industry component, the other substituting the between-industry component. These substitutions were made with respect to the equations in Table 4.

Since the changes induced in the parameter estimates for other causal variables were trivial, only the estimates for alternative forms of EMPCON are presented in Table 5. These results demonstrate that removing industry variation from the definition of EMPCON does not attenuate its association with strike activity. Both standardized and unstandardized coefficients estimated for the within-industry component do not deviate noticeably from their counterparts in Table 4. It is the between-industry component which exhibits the weaker relation with strike activity. However, particularly in the case of

Table 5. Coefficients for Within- and Between-Industry Components of EMPCON in Strike Measure Regressions *

Dependent Variables	Component of EMPCON	
	Within-Industry	Between-Industry
STRIKES	(.257) 18.502 (8.066)	(.055) 9.299 (20.056)
STRIKERS	(.240) 20.326 (7.673)	(.061) 12.292 (19.509)
MANDAYS	(.251) 23.241 (6.895)	(.093) 20.497 (18.407)
\bar{X}	0.754	0.752
S.D.	0.019	0.008

* These values were obtained by substituting each component for the original EMPCON measure in the Table 4 equations and reestimating them. The values presented are metric regression coefficients, parenthetical values being standardized regression coefficients (above) and standard errors (below).

¹⁷ This procedure may be summarized as follows: if X_{ij} is the CBP matrix of employing units by industry ($i=1, \dots, 37$) and size ($j=1, \dots, 8$) percentaged over the i rows, Y_{ij} is the same matrix aggregated over all SMSAs, and $X_{i.}$, $X_{.j}$ and $Y_{i.}$, $Y_{.j}$ are marginal arrays expressed as percentages of the SMSA and national totals, respectively; then it can be shown (although the proof is omitted here; see Kitagawa, 1955) that: $X_{i.} - X_{.j} = Y_{i.}(X_{.j} - Y_{.j}) + Y_{.j}(X_{i.} - Y_{i.}) + (X_{i.} - Y_{i.})(X_{.j} - Y_{.j})$. In other words, the percentage difference between the employment-size array for the SMSA and that for the nation is partitioned into three parts. The first is the difference that would obtain if the SMSA had the nation's industry distribution; the second is the difference that would obtain if the SMSA had the same industry-specific size distribution as the nation but a different industry distribution; the third is a covariance or interaction term made up of the industry difference and the industry-specific size difference. To eliminate negative signs, $Y_{.j}$ was added to each component distribution. I am indebted to Hal Winsborough of the University of Wisconsin, Madison, for suggesting this approach.

Table 6. Means and Standard Deviations by Region

Variable	South (N=21 SMSAs)		Non-South (N=57 SMSAs)	
	\bar{X}	SD	\bar{X}	SD
INDSP	0.090	0.010	0.094	0.020
ABSOWN	0.083	0.054	0.121	0.071
EMPCON	0.745	0.019	0.755	0.020
UNIZ	0.451	0.159	0.721	0.150
UCON	0.465	0.091	0.526	0.082
STRIKES *	4.447	1.595	5.216	1.190
STRIKERS *	5.495	1.710	6.435	1.465
MANDAYS *	8.079	2.108	9.109	1.494
SIZE *	12.014	0.646	12.449	1.124
MANU	0.165	0.062	0.288	0.107
NWHITE	0.211	0.105	0.071	0.066
UNEMP	0.043	0.012	0.052	0.014
EXPSTK (strikes) *	4.423	0.651	4.993	1.081
EXPSTK (strikers) *	5.833	0.696	6.457	1.127
EXPSTK (mandays) *	8.537	0.692	9.221	1.135

* Variable has been transformed to its natural logarithm.

MANDAYS, this apparent reduction is due mainly to the smaller variance of the between-industry component. It would appear that most of the size variation in employing organizations in these SMSAs lies within rather than between the industries defined by this classification. Given this additional evidence, it seems fair to say that the patterns observed in these data are not attributable to the industry composition of SMSAs.

Regional Patterns

A final source of variability in local strike activity is region. The South, especially, is a social and economic context within the United States which may require separate attention. As Table 6 shows, South/non-South differences exist in the means and standard deviations of the variables considered in this report.¹⁸ The well-known failure of the labor movement to make inroads in this region comparable to its achievements in other parts of the country (Roy, 1965) shows up clearly in regard to unionization (UNIZ). Among the 21 Southern SMSAs in this

sample, an estimated 45% of the total plant workers were under union contract in 1965, whereas 72% of such employment was unionized in the 57 metropolitan areas outside the South. Similarly, all three strike indicators show higher levels in non-Southern than in Southern SMSAs. But note the larger variances in the South.

There is also the possibility, however, that equations estimated for the total sample of 78 SMSAs obscure meaningful South/non-South differences in the kinds of causal processes generating union organization or industrial conflict within regions. Indeed, an important theme in historical and comparative research on strikes is that the institutional setting of industrial relations may specify the exact set of conditions which are causally antecedent to strikes (Ingham, 1974; Ross and Hartman, 1960; Snyder, 1975). To allow for this possibility, tests for interaction between region and the determinants under investigation here were made with regard to the union and strike measures. The results are presented in Table 7. In fact, there is no evidence that the within-region slopes differ by more than sampling error. None of the F-ratios for the R^2 increments due to interaction is significant. Nor are main effects of region present in four of the five equations. These results imply that the differences on the endogenous variables evident in Table 6 are attributable to the equally evident dif-

¹⁸ Twenty-one SMSAs in the following states were coded as present in the South: Texas, Oklahoma, Arkansas, Louisiana, Alabama, Mississippi, Florida, Georgia, North Carolina, South Carolina, Virginia, West Virginia, Kentucky, and Tennessee.

Table 7. Summary of Tests for Additive and Interaction Effects of Region on Union Organization and Strike Activity (N=78 SMSAs)

	UNIZ	UCON	STRIKES	STRIKERS	MANDAYS
R^2 for equation ^a					
1	.469	.592	.529	.698	.803
2	.573	.511	.534	.704	.809
3	.634	.567	.593	.738	.839
F-ratios ^b					
$R^2_3 - R^2_1$	16.830***	2.631	.772	1.364	2.019
$R^2_2 - R^2_1$	1.475	1.144	1.076	.980	1.978

^a With respect to UNIZ and UCON, Equation 1 is the same as in Table 2. With respect to the strike measures, Equation 1 is the same as in Table 4. Equation 2 differs from Equation 1 in including as a regressor SOUTH, a dummy variable for presence in the South. Equation 3 differs from Equation 2 in including interaction terms formed as the products of SOUTH with the other regressors in these equations.

^b One, two, and three asterisks indicate significance levels of .05, .01, and .001, respectively.

ferences on the regressors. In the case of UNIZ, however, the South/non-South discrepancy remains with these factors held constant. Indeed, controlling the covariates reduces the regional discrepancy in unionization only slightly from 27% to 23%; this is the appropriate interpretation for the metric regression coefficient on SOUTH (a dummy variable scored one for SMSA presence in that region, zero for absence) in the following equation:

UNIZ =

$$\begin{aligned}
 & 0.203 + 0.110\text{INDSP} + 0.471\text{ABSOWN} \\
 & \quad (.001) \quad (.167) \\
 & \quad (1.147) \quad (0.304) \\
 & (-.505) \quad (.225) \quad (.250) \\
 & -0.475\text{EMPCON} + 0.042\text{SIZE} + 0.434\text{MANU} \\
 & \quad (1.264) \quad (0.020) \quad (0.256) \\
 & (.225) \quad (2.40) \quad (.552) \\
 & + 2.575\text{UNEMP} + 0.465\text{NWHITE} - 0.226\text{SOUTH} \\
 & \quad (1.152) \quad (0.211) \quad (0.055)
 \end{aligned}$$

$$R^2 = .573 \quad \bar{R}^2 = .531.$$

The parenthetical values above the metric regression coefficients are standardized regression coefficients; those below are standard errors. Not only does SOUTH emerge as the best predictor of this measure of labor force unionization, but its inclusion in the equation produces some notable changes in the coefficients paired with other regressors (see Table 2). In particular, the effects of ABSOWN and MANU are attenuated now that region is controlled. Note also that a positive significant coefficient is now associated with

NWHITE. Contrary to Seidman's (1965) hypothesis concerning the effect of ethnic diversity on industrial relations, it appears that, within regions, an increasingly non-white labor force implies greater unionization (see Hill, 1974, and Leggett, 1968).

CONCLUSIONS

Because they saw absentee ownership and industrial specialization giving rise to homogeneous concentrations of workers geographically removed from their employers, Kerr and Siegel argued that these two properties of community structure should produce a strike-prone labor force. Moreover, they saw the strikes arising from such conditions as spontaneous, rank-and-file actions, not contingent on the behavior of unions. The findings produced in the present study, however, indicate that absentee ownership and industrial specialization *only* influence metropolitan strike rates insofar as they contribute to the level and form of unionization in the SMSA. That is, given a model that allows for both direct effects of community structure on strike activity as well as indirect effects through the intervening process of union organization, the data considered here reveal evidence only of indirect effects.

Moreover, it is uncertain whether absentee ownership and industrial specialization affect urban unionism through the processes conceived by Kerr and Siegel or even Marx: the aggregation of

discontent produced by workers having high rates of contact with one another while facing general segregation from the broader community. Consider the measured effect of industrial specialization on union concentration. Insofar as unions organize workers on an industry-wide basis, one would expect to find fewer but larger unions in industrially homogeneous communities. A wide variety of industries in an area, each employing a relatively small fraction of the labor force, should therefore result in an equally fragmented labor movement. The finding that industrial specialization affects only union concentration and has nothing to do with unionization is compatible with this interpretation but is not easily accounted for within the isolated mass framework.

As suggested earlier, a similarly competing interpretation exists for the association found between absentee ownership and unionization. Several studies indicate that a high degree of absentee ownership contributes to unionization (if not strikes) by shifting community support from the side of management to that of workers in industrial disputes. The coefficient relating absentee ownership to unionization in this analysis is, in any case, quite weak, although one should not discount the possibility that it is attenuated because of the sources of error noted previously.

But this pessimistic appraisal of the principles advanced by Kerr and Siegel for explaining patterns of strike activity must be revised when one considers the case of employment concentration. If industrial specialization increases the homogeneity of the work force, employment concentration defines the actual organizational boundaries within which interaction networks among workers are formed. In one sense, then, it is a better indicator of the degree to which employees are actually placed in contact with one another in a particular urban setting. Two findings of this investigation would appear to offer rather strong support for the kind of argument formulated by Kerr and Siegel. First, the marked effects of employment concentration on strike activity are entirely direct, wholly unmediated by the character of local unionization. Sec-

ondly, such effects become stronger as the unionization of the labor force and the centralization of union power declines. This last is a pattern which further supports Kerr and Siegel's assertion that unions are not essential for strikes to result from isolated mass conditions.

Still, the finding that the combination of strong, centralized unions and large employing organizations depresses metropolitan strike activity over and above the effects of these variables acting alone may also be explained by another perspective (see Ross and Hartman, 1960, and Ingham, 1974). Several observers have viewed the continuing concentration in the organizational power of corporations and unions as a process creating a pattern of collective bargaining in which strikes are symbolic and infrequent. Industrial relations under these conditions become ritualized, predictable affairs with only a narrow range of issues negotiated between the elite representatives of corporate and labor bureaucracies. The tendency observed here for SMSAs with highly centralized employment and union structures to have lower strike rates may well be explained as the outcome of processes such as these.

Perhaps more important than the question of whether they support the Kerr-Siegel theory is simply that these data indicate rather clearly that the manner in which employers and workers are organized in urban communities has distinct consequences for local levels of industrial conflict. Having discussed at some length the implications of such aggregate properties of local employment structure as absentee ownership, industrial specialization, and employment concentration, a final comment regarding the observed effects of the union variables seems appropriate. Given the recent interest in the shape of strikes (e.g., Britt and Galle, 1974; Shorter and Tilly, 1971), a particularly important finding here is that pertaining to the distinctive pattern of effects displayed by UNIZ and UCON. If sheer membership in unions is associated with the frequency of strikes in SMSAs, it is much less important as a determinant of their form (i.e., size and duration) than the degree to which the administrative re-

sources possessed by all labor organizations in the community are concentrated in a few large unions. Although the strength of these effects has been shown to vary with the level of employment concentration, there is still a fairly clear tendency for strikes to be larger and longer in SMSAs where union staff are highly centralized. Thus, even within the rather rigid constraints set by institutional unionism in the contemporary United States, local variations in the organization of workers viz-à-viz employers retain a significant effect on the pattern of urban industrial conflict.

REFERENCES

- Aiken, Michael
1970 "The distribution of community power: structural bases and social consequences." Pp. 487-525 in Michael Aiken and Paul E. Mott (eds.), *The Structure of Community Power*. New York: Random House.
- Aiken, M. and Robert R. Alford
1970 "Community structure and innovation: the case of urban renewal." *American Sociological Review* 35:650-65.
- Alford, Robert R.
1963 *Party and Society*. Chicago: Rand McNally.
- Amsden, Jon and Stephen Brier
1977 "Coal miners on strike: the transformation of strike demands and the formation of the national union in the U. S. coal industry, 1881-1894." *Journal of Interdisciplinary History* 7:583-616.
- Ashenfelter, Orley and George E. Johnson
1969 "Bargaining theory, trade unions, and industrial strike activity." *American Economic Review* 59:35-49.
- Ashenfelter, Orley and John H. Pencavel
1969 "American trade union growth, 1900-1960." *Quarterly Journal of Economics* 83:434-48.
- Bendix, Reinhard and Seymour Martin Lipset
1966 "Karl Marx's theory of social classes." Pp. 6-11 in Reinhard Bendix and S. M. Lipset (eds.), *Class, Status, and Power*. 2nd ed. New York: Free Press.
- Blauner, Robert
1964 *Alienation and Freedom: The Factory Worker and His Industry*. Chicago: University of Chicago Press.
- Bogue, Donald J.
1969 *Principles of Demography*. New York: Wiley.
- Britt, David and Omer Galle
1972 "Industrial conflict and unionization." *American Sociological Review* 37:46-57.
1974 "Structural antecedents of the shape of strikes: a comparative analysis." *American Sociological Review* 39:642-51.
- Cass, Millard
1957 "The relationship of size of firm and strike activity." *Monthly Labor Review* 80: 1330-4.
- Dahrendorf, Ralf
1959 *Class and Class Conflict in Industrial Society*. Stanford: Stanford University Press.
- Derber, Milton, W. E. Chalmers, and R. Stagner
1958 "Environmental variables and union-management accommodation." *Industrial and Labor Relations Review* 11:413-28.
- Duncan, O. D.
1970 "Partials, partitions, and paths." Pp. 38-47 in E. F. Borgatta and G. W. Bohrenstedt (eds.), *Sociological Methodology*. San Francisco: Jossey-Bass.
- Fortune
1963 *Plant and Product Directory*. Vol. 1. New York: Time
- Fiese, C. Frederick
1974 "Organization size, technology, and frequency of strikes." *Industrial and Labor Relations Review* 27:560-71.
- Fischer, Claude
1973 "On urban alienations and anomie: powerlessness and social isolation." *American Sociological Review* 38:311-26.
- Fuguitt, Glenn V. and Stanley Lieberman
1874 "Correlations of ratios or difference scores having common terms." Pp. 128-44 in H. L. Costner (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.
- Gibbs, Jack P. and Walter T. Martin
1962 "Urbanization, technology, and the division of labor: international patterns." *American Sociological Review* 27:667-77.
- Gibbs, Jack P. and Dudley L. Poston, Jr.
1975 "The division of labor: conceptualization and related measures." *Social Forces* 53:468-75.
- Goldberg, Albert I.
1974 "The local community and teacher strike solidarity." *Industrial Relations* 13:288-98.
- Griffin, John I.
1939 *Strikes: A Study in Quantitative Economics*. New York: Columbia University Press.
- Hawley, Amos H.
1950 *Human Ecology: A Theory of Community Structure*. New York: Ronald.
- Hill, Richard Child
1974 "Unionization and racial income inequality in the metropolis." *American Sociological Review* 39:507-22.
- Ingham, Geoffrey K.
1970 *Size of Industrial Organization and Worker Behavior*. London: Cambridge University Press.
1974 *Strikes and Industrial Conflict: Britain and Scandinavia*. London: Macmillan.
- Eerr, Clark and Abraham Siegel
1954 "The interindustry propensity to strike." Pp. 189-211 in Arthur Kornhauser, R. Dubin, and A. Ross (eds.), *Industrial Conflict*. New York: McGraw-Hill.
- Kitagawa, Evelyn M.
1955 "Components of a difference between two

- rates." *Journal of the American Statistical Association* 50:1168-95.
- Knowles, K. G. J. C.
1952 *Strikes: A Study in Industrial Conflict*. Oxford: Blackwell.
- Kornblum, William
1974 *Blue Collar Community*. Chicago: University of Chicago Press.
- Leggett, John
1968 *Class, Race, and Labor*. New York: Oxford University Press.
- Lieberson, Stanley
1969 "Measuring population diversity." *American Sociological Review* 34:850-62.
- Lincoln, James R.
1976 "Power and mobilization in the urban community: reconsidering the ecological approach." *American Sociological Review* 41:1-15.
- Lipset, Seymour Martin
1963 *Political Man*. New York: Doubleday.
- Lynd, Robert S. and Hellen M. Lynd
1937 *Middletown in Transition*. New York: Harcourt Brace.
- Marx, Karl
[1963] "The eighteenth Brumaire of Louis Bonaparte." Pp. 318-48 in L. S. Feuer (ed.), *Marx and Engels: Basic Writings on Politics and Philosophy*. New York: International.
- Mills, C. W. and M. J. Ulmer
1970 "Small business and civic welfare." Pp. 124-53 in M. Aiken and P. E. Mott (eds.), *The Structure of Community Power*. New York: Random House.
- O'Brien, F. S.
1965 "Industrial conflict and business fluctuations: a comment." *Journal of Political Economy* 73:650-4.
- Olson, Mancur, Jr.
1971 *The Logic of Collective Action*. New York: Schocken.
- Pellegrin, R. J. and C. H. Coates
1956 "Absentee-owned corporations and community power structure." *American Journal of Sociology* 61:413-39.
- Pencavel, John H.
1970 "An investigation into industrial strike activity in Britain." *Economica* 37:239-56.
- Pope, Liston
1942 *Millhands and Preachers: A Study of Gastonia*. New Haven: Yale University Press.
- Rees, Albert
1954 "Industrial conflict and business fluctuations." Pp. 213-20 in Arthur Kornhauser, R. Dubin, and A. Ross (eds.), *Industrial Conflict*. New York: McGraw-Hill.
- Revens, R. W.
1956 "Industrial morale and size of unit." *Political Quarterly* 27:303-11.
- Ross, Arthur M. and Paul T. Hartman
1960 *Changing Patterns of Industrial Conflict*. New York: Wiley.
- Roy, Donald F.
1965 "Change and resistance to change in the Southern labor movement." Pp. 225-47 in J. C. McKinney and E. T. Thompson (eds.), *The South in Continuity and Change*. Durham: Duke University Press.
- Schulze, Robert O.
1958 "The role of economic dominants in community power structure." *American Sociological Review* 23:3-9.
- Seidman, Joel
1965 "Community influences on industrial relations." *Industrial Relations* 2:1-7.
- Shorter, Edward and Charles Tilly
1971 "The shape of strikes in France, 1830-1960." *Comparative Studies in Society and History* 13:60-86.
- 1974 *Strikes in France, 1830-1968*. London: Cambridge University Press.
- Skeels, Jack W.
1971 "Measures of U. S. strike activity." *Industrial and Labor Relations Review* 24:515-25.
- Snyder, David
1975 "Institutional setting and industrial conflict: comparative analyses of France, Italy, and the United States." *American Sociological Review* 40:259-78.
- Snyder, D. and William R. Kelley
1976 "Industrial violence in Italy, 1878-1903." *American Journal of Sociology* 82:131-62.
- Spilerman, Seymour
1970 "The causes of racial disturbances: a comparison of alternative explanations." *American Sociological Review* 35:627-49.
- 1971 "The causes of racial disturbances: tests of an explanation." *American Sociological Review* 36:427-42.
- 1976 "Structural characteristics of cities and the severity of racial disorders." *American Sociological Review* 41:771-93.
- Stearns, Peter N.
1974 "Measuring the evolution of strike movements." *International Review of Social History* 19:1-27.
- Stern, Robert N.
1976 "Intermetropolitan patterns of strike frequency." *Industrial and Labor Relations Review* 29:218-35.
- Stinchcombe, Arthur L.
1965 "Social structure and organizations." Pp. 142-93 in James G. March (ed.), *Handbook of Organizations*. Chicago: Rand McNally.
- Thompson, Wilbur R.
1965 *A Preface to Urban Economics*. Baltimore: Johns Hopkins University Press.
- U.S. Bureau of the Census
1960 *Characteristics of the Population*. Washington, D.C.: U.S. Government Printing Office.
- 1967 *County Business Patterns*. Washington, D.C.: U. S. Government Printing Office.
- U. S. Bureau of Employment Security, Department of Labor
1963- *Area Labor Market Trends*. Washington, D.C.: U. S. Employment Service.
- U.S. Bureau of Labor Statistics
1966 *Wages and Related Benefits: Part I, Metropolitan Areas*. Bulletin No. 1465-86. Washington, D.C.: U. S. Government Printing Office.

- 1964- Analysis of Work Stoppages. Washington, D.C.: U. S. Government Printing Office.
- Warner, W. Lloyd and J. O. Low
1947 The Social System of the Modern Factory. New Haven: Yale University Press.
- Weintraub, Andrew
1966 "Prosperity vs. strikes: an empirical approach." *Industrial and Labor Relations Review* 19:231-8.
- Wray, Donald E.
1953 "The community and labor-management relations." Pp. 1-45 in W. E. Chalmers (ed.), *Labor-Management Relations in Illini City*. Urbana: University of Illinois Institute of Labor and Industrial Relations.
- Food and Kindred Products
Textile Mill Products
Apparel
Lumber, Wood Products, Furniture
Printing and Publishing
Chemicals and Allied Products
Electrical and Other Machinery
Motor Vehicles and Equipment
Other Transportation Equipment
Other and Miscellaneous Manufacturing
Railroads and Railway Express
Trucking and Warehousing
Other Transportation
Communications
Utilities and Sanitary Service
Wholesale Trade
Retail Trade
Finance, Insurance, and Real Estate
Hotels and Other Personal Services
Business and Repair Services
Entertainment and Recreation Services
Medical and Other Professional
Public Administration

APPENDIX

26-Item Industry Classification

Agriculture, Forestry, Fisheries
Mining
Contract Construction

STATUS GENERALIZATION: A REVIEW AND SOME NEW DATA*

MURRAY WEBSTER, JR.
JAMES E. DRISKELL, JR.
University of South Carolina

American Sociological Review 1978, Vol. 43 (April):220-236

Theoretical developments and empirical evidence over the past two and a half decades are reviewed with an attempt to assess the state of sociological knowledge about how, and conditions under which, external status characteristics of individuals affect their face-to-face interaction. This review describes the development of expectation states theory, which assumes a burden of proof process by which initially unrelated status characteristics become relevant to task performance, and task specific expectations determine patterns of interaction. Three versions of the theory, with increasing scope, are described and evidence relevant to them is summarized. Data from an independent test of the most recent version are presented, and the review concludes with a summary of knowledge and some suggestions for applications of the theory of status generalization effects.

Status generalization, the process by which statuses of actors external to a particular interaction are imported and allowed to determine important features of that interaction, has been studied by sociologists since the early days of our discipline. By the 1950s the following

generalizations were fairly well accepted by most investigators in the field:¹

1. Status characteristics such as race, age, sex, and occupation act as cues to individuals and are used to order their interactions with persons previously unknown to them.

* This report was prepared during a summer research leave to Murray Webster. New experimental data were collected as part of James Driskell's M.A. thesis, and Leroy F. Smith and Donia Gardner helped as coexperimenters. Financial support was provided by Vice-President H. Willard Davis and Provost Keith E. Davis.

¹ The volume of relevant literature precludes an adequate list of citations. Early theorists in this area include Simmel (1908), Park (1925; Park and Burgess, 1921), and Hughes (1945). For a good summary of the effects of race, see Katz (1970); for sex, Meeker and O'Neill (1977); for occupation, Berger et al. (1972).

2. The significant feature of status characteristics is that they are culturally *evaluated*: it is preferable (socially desirable, invidious) to be white, adult, male, and managerial or professional.

3. The significant interaction features affected by status characteristics all involve subordination and superordination; such behaviors as deference, choices for group leader, estimation of individuals' contributions to group problem solution, being given and accepting chances to participate in discussions, exerting influence over others, being granted respect or esteem, and *proacting* in task oriented categories of interaction rather than *reacting* in social-emotional categories.

4. External evaluation of a status characteristic is maintained when it is imported. Thus high status outside the group is converted into high status within the group, and vice versa.

5. The status characteristic need not be relevant to the immediate interaction to order it; thus blacks, females, or manual laborers are treated as inferior to whites, males, or managers in group situations where no knowledge exists as to whether race, sex, or occupation differences are important in fulfilling the group task.

6. Status generalization is a process—not always conscious—used by individuals to structure unfamiliar social situations. This implies (a) that this is the crucial process which permits interaction to occur; and (b) that status generalization is most effective in ordering interaction where individuals have no prior interaction history, no information about one another except for their known statuses, and no general information about or experience with their task.

Within the past 25 years, theoretical and empirical research has confirmed the general outlines of these ideas. There have been, however, significant modifications and scope limitations placed upon them, increasingly precise statements of some, and significant additions to our total information about status generalization processes. Our purposes are (1) to review

these developments from the perspective of a theory of status characteristics and social interaction, (2) to present results of an experiment conducted to test the latest version of that theory, and (3) to identify some remaining areas of indeterminacy in the knowledge. A series of theoretical treatments by Berger et al. (1966; 1972; 1974; 1977) have consistently dealt with the phenomenon of status generalization. Our presentation uses terms and ideas from that tradition.²

Berger's work is concerned with two general processes: social conditions affecting individuals' formation of *performance expectations*, and the effects of those expectations on subsequent interaction in the group. Prior to 1966, theoretical and empirical work had shown that individuals who begin interaction as status equals (such as the white, male, high SES college students of Bales groups) will develop expectations—roughly equivalent to ability conceptions—for each others' performances as the result of problem-solving interaction. The structure of expectation states in the group then will determine all observable features of interaction inequality. The higher the expectations associated with a given actor, (1) the more likely he is to receive and to accept chances to perform; (2) the more likely he is to receive agreement and other types of esteem from other members; (3) the more likely his performances are to be evaluated positively; (4) the more likely he is to be influential and the less likely he is to accept influence from others in case of disagreement; and (5) the more likely he is to receive votes as "most helpful," or as "group leader." All of these interaction components tend to be highly intercorrelated, and together they constitute the observable power and prestige order of the group (Berger and Conner, 1969; Berger et al., 1969). Individuals come to be treated unequally *because* they are thought to

² Joseph Berger, Morris Zelditch, Jr., Rita J. Simon, and others gave helpful suggestions for improving earlier versions of this paper. We thank them, and accept responsibility for cases where we did not follow their advice. The content and organization of this review, the interpretation of results, and the emphasis given to various findings are ours.

have unequal task-relevant abilities in these groups.³

Once this theoretical work appeared, it became apparent that these interaction inequalities among individuals of equal external status, for whom unequal performance expectations are held, are similar to those described earlier among group members of differing status outside the group. That is, the interaction disability faced by blacks, females, or children has the features we would expect *if* we knew that these individuals had differing performance expectations associated with them—low expectations for individuals with low external status, and high expectations for individuals with high external status. It is as though external status differentiation functions to determine actors' expectations and behavior from the very beginning of interaction.

At the same time there was abundant data from the developing literature in small groups to show that, in a variety of dissimilar settings, groups developed a power and prestige structure consistent with members' external status positions: in Air Force bomber crews (Torrance, 1954); psychiatric hospital teams (Caudill, 1958); juries (Strodtbeck and Mann, 1956; Strodtbeck et al., 1958); and boys at summer camps (Sherif et al., 1955). Most of this research was viewed as descriptive and possibly pretheoretical, as illustrated by this concluding statement from the well-known study by Strodtbeck et al. (1958:388) on jury deliberations: "While our data do little to illuminate *how* differentiation arises, they show that status gradients clearly emerge. . . . [T]his demonstration of status continuity should be noted in any theory concerned with describing the process of status affirmation and maintenance."

These cases and many others were explained by the first theory of status characteristics and expectation states (Berger et al., 1966). According to this theory, interaction studies involving such characteristics as race, sex, occupation, and

military rank are related in that they all involve *status characteristics*. Status characteristics operate to define interaction situations through a generalization principle called "burden of proof," and determine group behavior through an expectation state process. The burden of proof principle states that, unless the relevance or applicability of an external status characteristic is challenged, actors will *infer* task-specific performance expectations on the basis of any discriminable status characteristics they possess. Such inferences occur *regardless of the actual relevance* of the status characteristics to the task at hand. In other words, individuals act as if the burden of proof is placed upon demonstrating that the status characteristic is *not* relevant rather than demonstrating that it *is* relevant to the task at hand. In the absence of such demonstration, individuals will treat the status characteristic as relevant, and status generalization will occur.

To state this idea more precisely; we introduce some theoretical notation to be used throughout this paper.

All versions of the theory apply to groups of individuals (p and o) meeting two scope conditions, task orientation and collective orientation. Task orientation means that the members' purpose in meeting is to solve some problem, rather than simply to enjoy each other's company. Collective orientation means that they consider it legitimate and necessary to consider every individual's ideas in working on the task. Any situation meeting these conditions will be called "task situation S," or S. Let D represent an external diffuse status characteristic such as age, sex, race, or SES; and let C* represent the ability characteristic necessary to complete the group's task.

Then:

Assumption 1. (Activation) D is activated in S if D is a basis of discrimination between p and o.

Assumption 2. (Burden of Proof) If C* has not been previously dissociated from D, and if D is activated in S, and D is the only basis discriminating p and o,

³ A summary of this work is Webster, 1975: Chap. 6. Notice that the meaning of expectations here is different from that used by some others; among them Archibald, 1976; McCranie and Kimberly, 1973; and Sampson, 1963.

then D will become relevant to C* in S.

Assumption 3. (Assignment) If D is relevant to C* in S, p will assign states of C* that are consistent with the states of D which p and o possess.

Assumption 4. (Basic Expectation Assumption) If p assigns states of C* to himself and o, then p's position relative to o in the observable power and prestige order will be a direct function of p's expectation advantage over o (adapted from Berger et al., 1966:29-39; and Berger et al., 1972:245-7).

Less formally, assumption 1 states that any D which discriminates p and o will become salient to them in forming expectations. Assumptions 2 and 3 state that when actors are differentiated by a single D, unless they know for sure that the D is irrelevant to C*—as they might in an experiment in which the experimenter told them that, say, sex (D) was irrelevant to jury deliberation skill (C*)—they will treat it as though it is relevant, and form task specific expectations consistent with it. Assumption 4 says that all features of interaction inequality (including receiving and accepting chances to perform, positive and negative performance evaluations, agreement, esteem, authority, and acceptance or rejection of influence) depend on the relative expectations associated with group members. Thus a feature of a culture—that status characteristics carry both evaluational and performance connotations—is linked to expectations through a social psychological process (burden of proof) and affects structural features of problem solving groups (their observable interaction inequality structures).

The burden of proof idea was tested directly and confirmed in two experiments (Moore, 1968; Berger et al., 1972) designed for that purpose.⁴ The first used the

status characteristic academic attainment, and the second used Air Force rank. In some conditions of each experiment the status characteristic was made directly relevant to performance at the group's task, and in other conditions relevance was not mentioned. Status generalization occurred in all cases.⁵

So from this first version of the theory and these two experiments, there was a confirmed theoretical explanation for the status generalization process in at least one type of situation, that involving a single status characteristic. The explicit propositions of the 1966 theory specified conditions under which status generalization would occur (situation S), how it occurred (burden of proof), what the specific effect of the status difference was (to determine performance expectations), and how status ordered interaction (as expectation states determined the observable power and prestige order of groups).

Multiple Characteristics

The next theoretical question is, what happens when two or more status char-

variants of the basic expectation experiment, a standardized situation used to evaluate theoretical developments. The basic expectation experiment has two parts, or phases. In phase 1, all status information is introduced, and in phase 2 effects of the status information upon expectation states are measured. The usual number of subjects in each group is two, comparable to p and o of the assumptions above. Status information, such as differences in sex, age, race, or occupation of subjects, is presented by the experimenter (or, in later experiments, subjects are shown fictitious partners on closed circuit TV). Expectations are measured by one of the behaviors predicted to be affected by them: the probability of resolving experimentally introduced disagreements in favor of self, or P(s). Please see Webster and Sobieszek (1974: App. 1) for more details of the basic expectation experiment.

⁵ Berger et al. (1972) used degree of task relevance as a second independent variable. When rank only was mentioned, the P(s) difference between high D and low D conditions was .06; when rank was mentioned and said to be associated with a variety of skills, but not including the phase 2 task the difference was .08; when rank was mentioned, linked to other tasks and also to the criterion task, the difference was .14. But the 1966/72 theory did not predict any difference in effect by degree of relevance, so the authors did not interpret these differences. This differential effect is incorporated in the third version of the theory, discussed below.

⁴ Many experimental tests of the theory rely on

acteristics are salient in an interaction situation? Such cases include, for instance, Hughes's (1945) classic dilemma (for the male patient) of seeking medical treatment from a female physician. Multiple characteristic situations actually contain two major theoretical issues: first, the necessity of differentiating between types of status characteristics; and second, the problem of describing how these different characteristics are processed to form aggregate performance expectations.

When we speak of a woman doctor, we speak of two different types of statuses. The important question is, what makes them different, especially since we have already noted above that both sex (Strodtbeck and Mann, 1956; Meeker and O'Neill, 1977) and occupation (Strodtbeck et al., 1958; Berger et al., 1972) produce task specific expectations through a burden of proof process. Woman is the low evaluated state of sex (in most cultures), and doctor is a positively evaluated state of occupation (in our culture), but the two characteristics differ in scope, with sex being much broader. Occupation produces only a limited set of expectations, and only in a limited set of social situations. One expects a physician to be good at healing and comforting in a clinic, but not necessarily good at many other tasks such as athletics or repairing machinery. By contrast, sex is more diffuse; it carries all sorts of connotations of capacity and incapacity. In our culture, males are widely thought to be more intelligent, logical, sensible, mechanical, athletic, strong, etc., than females. The etc. is important: doctor carries expectations of delimited scope; sex carries expectations without (explicit) limit. The same is true for the specific characteristic reading ability, which carries expectations only in certain specifiable circumstances; and for the diffuse characteristic race which affects expectations in a much wider range of situations.

To capture this difference between types of status characteristics, two types, D and C, were identified. Type C refers to statuses such as occupation, skill, or training. More formally, a characteristic C is a *specific status characteristic* if:

1. the states of C are differentially evaluated in the culture;
2. to each state (+ or -) of C there corresponds a distinct, similarly evaluated expectation state having the same evaluation as C and relevant to a particular task.

The second type, D, refers to statuses such as race, sex, and age; those describing global properties of an actor. More formally, a characteristic D is a *diffuse status characteristic* if:

1. the states of D are differentially evaluated in the culture;
2. to each state (+ or -) of D there corresponds a set of similarly evaluated states of specific, evaluated characteristics associated with D;
3. to each state of D there corresponds a similarly evaluated general expectation state (GES) having the same evaluation as D (adapted from Berger et al., 1974:175).

Thus physician is a specific status characteristic (C) if (1) it is considered high prestige to be a physician and low to be a waitress; and (2) if physicians are expected to be more competent than waitresses at healing. Similarly, sex is a diffuse status characteristic (D) if (1) it is considered preferable to be male (or female) and (2) males (or females) are assumed to be more intelligent and better at most tasks than females (or males). Note that most specific occupations probably fit the definition of a C, while SES usually is a D because it carries a global set of ideas of competences, or a GES. Below it will be important to distinguish between D's and C's, as of course it is for many specific applications of the theory. However what is significant at the moment is that the burden of proof status generalization process operates for C's just as it does for D's (Freese, 1970; Kevin, 1975).

The second issue, how multiple items of status information are processed, may be more difficult. Essentially there are two ways in which information on multiple status characteristics could be used. Individuals might simplify the situation cognitively by choosing only one characteristic

and using it to order their interactions. Let us call this process "elimination," to indicate that all available characteristics except one will be eliminated from consideration in expectation formation. The elimination process is represented by a complaint such as "No matter how much I know about the problem, and no matter what my credentials, they don't listen to my ideas because I'm a (woman, black, child)."⁶ Or, individuals might use status information in what we will call a "combining process," taking in all positively and negatively valued characteristics and combining them to form intermediate expectations. The combining process is represented by statements such as "She may be a doctor, but she's also a woman (so she's not quite as capable as a male doctor)."⁷

The fact that burden of proof is known to work for both D's and C's is important here because the first information on combining vs. eliminating comes from experiments in which subjects are differentiated by C's, not D's. Three laboratory experiments were conducted to explore directly whether formation of aggregate expectations could better be represented by a combining or an elimination process. All of these used C's, manipulated by administering purported tests of various abilities, as status information differentiating subjects.

In the first experiment (Berger and Fisek, 1970), male college student subjects were told they were either high on both C_1 and C_2 , high on C_1 and low on C_2 ,

or low on C_1 and C_2 . Aggregate expectations as measured by the standard $P(s)$ data, were ordered as follows: $1 > 2 > 3$. This ordering (and the pattern of variance within conditions) is consistent only with a combining process.

The second of these exploratory experiments (Webster et al., 1972), showed essentially the same result with male and female college student subjects. The experiment provided individuals status information on a single C, but instead of getting definite information from the experimenter, subjects received evaluative opinions from two other individuals, potential significant others, who sometimes agreed and sometimes disagreed. Disagreement among evaluators, as well as some degree of uncertainty over how accurate their opinions were, may have increased pressure to employ an elimination process in this experiment. However, again the evidence showed good support for a simple combining process and no support for elimination.⁸

The third experiment was some additional groups run in the Berger-Fisek situation above (Berger and Fisek, 1974). In the first additional condition, subjects were high on both C_1 and C_2 , while partners were high on C_1 and low on C_2 . The second additional condition was the mirror image: subjects were high on C_1 and low on C_2 , while their partners were high on both C_1 and C_2 . These additional conditions were run to examine the possibility that the finding from the first three conditions might have been an artifact of the exact symmetry of status information in the earlier condition 2. But again evidence supports a combining, not an elimination process. If we call the additional conditions 4 and 5, then the complete ordering of all these conditions by $P(s)$ is: $1 > 4 > 2 > 5 > 3$, the ordering predicted by a combining hypothesis.

During this time, a second group of experiments was conducted to test two variant versions of expectation states theory (Freese, 1970; Kervin, 1972). These data

⁶ Much of the status consistency literature of the 1950s and 1960s assumed that individuals possessing inconsistent statuses will follow an elimination process, and furthermore, that they will try to make their own highest status the only one relevant to any interaction. Thus a black man interacting on a jury with a white woman would be expected to eliminate the race status in forming his expectations for self and other, and the woman would be expected to eliminate sex. So also, status conflicts with resulting psychological stress often were expected. See Lenski, 1954; 1956; 1966; Jackson, 1962; Jackson and Burke, 1965; Hornung, 1977.

⁷ Psychologists' work on impression formation and person perception, in which subjects are asked to form an overall impression of a hypothetical other after reading a list of adjectives purported to describe him, often finds support for one or another of several combining models. See Anderson, 1968; 1977.

⁸ These experiments compared several specific information processing models of combining and eliminating. Data support only the combining models, with simpler combining processes generally receiving stronger support than more complex.

also bear on the issue of combining vs. elimination. Freese (1970) used female subjects recruited from a temporary employment service. In this experiment only one characteristic was measured, but it was linked cognitively to a second, unmeasured characteristic. In condition 1, subjects were told they possessed the high state of C_1 , and that people high on C_1 were usually high also on a second, unmeasured characteristic, C_2 . In condition 2, the subject was low on C_1 , and people low on C_1 were said to be usually low on C_2 also. In condition 3, the subject was high on C_1 , but this was said to be associated with being low on C_2 ; and in condition 4, the subject was low on C_1 and this was associated with being high on C_2 . Our interpretation of results is that burden of proof functioned in all cases, with combining of both measured and inferred characteristics. $P(s)$ of condition 1 was .73, reflecting the high states of C_1 and C_2 ; and for condition 2 it was .62, reflecting the low states of the C 's. In conditions 3 and 4, the $P(s)$ was intermediate (.66 for both conditions), reflecting aggregate expectations formed from both of the inconsistent C 's subjects possessed. (For a somewhat different interpretation of this experiment, see Freese, 1976.)

Kervin (1972; data published in Kervin, 1977) conducted an experiment with male high school students as subjects. In his condition 2, subjects possessed the high states of C_1 and C_2 while their partners possessed the low states of both. In condition 3, subjects possessed only the high state of C_1 , and their partners the low state. In condition 4, subjects, compared to their partners, were high on C_1 and low on C_2 . Conditions were ordered $2 > 3 > 4$ in $P(s)$. Comparison of conditions 3 and 4 is most significant here, for it shows that the low state of C_2 was combined with the high state of C_1 in condition 4, and thus produced lower aggregate expectations here than in condition 3. Thus we conclude that both the Freese and the Kervin experiments show evidence for a combining process.

The third group of studies relevant to the issue of combining vs. elimination were conducted in naturalistic settings, for applied purposes. Sociologists' recog-

nition of interaction problems in mixed status interaction, and attempts to redress the problems, antedate theoretical formulation of the problem in terms of status characteristics and expectation states. One early program of study is that of Katz and associates (Katz et al., 1958; Katz and Benjamin, 1960; Katz and M. Cohen, 1962; Katz, 1970).

These investigators worked with biracial groups (black and white) of college age men from varying geographical and SES backgrounds. Early studies (Katz et al., 1958) showed that whites dominated group interaction, and this interaction pattern was largely unaffected by several interventions designed to promote interracial harmony. A later study (Katz and Benjamin, 1960) showed that neither intellectual ability (IQ) nor authoritarianism (F-scale) was a cause of the interaction inequality. Finally, Katz and M. Cohen (1962) tried giving black members assertion training. This did increase blacks' participation rates somewhat (though not up to the level of whites), but it also aroused considerable hostility and downgrading of blacks by whites.

To interpret these results, we begin by noting that race is a diffuse status characteristic, with white being the preferred state. Then race functions through the burden of proof process to produce low expectations for blacks (held by both blacks and whites). These expectation differences produce differential participation rates and influence, both favoring whites. The assertion training in the Katz and M. Cohen study may well have raised blacks' expectations for themselves, since the training involved positive evaluations of performance, which are the source of expectations in cases where burden of proof does not operate. But whites' expectations for blacks were unaffected by the assertion training, so what ensued was a status struggle comparable to that seen in some Bales groups.⁹ From the whites'

⁹ A status struggle, competition for discussion dominance, was once thought to be an inevitable feature of Bales groups. However Lewis (1972) showed that status struggles occur only under special conditions of complete initial status equality or when members focus on different status characteristics so that two or more of them think they have the highest

point of view, the blacks behaved inappropriately, considering their low abilities; from the blacks' point of view the whites, refusing to acknowledge the blacks' equal abilities, behaved inappropriately. Or less technically, the whites seemed like racists and the blacks seemed uppity.

E.G. Cohen's work (Cohen et al., 1970; Cohen, 1972; Cohen and Roper, 1972; Cohen et al., 1976) explicitly adopts the expectation theory perspective, and builds on Katz's work in several ways. These studies use male junior high school students in classroom-like situations. Early experiments (Cohen et al., 1970; Cohen, 1972) identified blacks' interracial interaction disability as an expectation problem along the lines described above, and concentrated on raising black children's self-expectations. But, as Katz found, this procedure did not produce status equal interaction, and it did increase the level of hostility in the groups.

Later experiments (Cohen and Roper, 1972; Cohen et al., 1976) treat expectations of both black and white boys, and have been successful at producing status-equal interaction.¹⁰ This work demonstrates the feasibility of overcoming the effects of race as a status characteristic, along lines implied by the theory.

Note we say the effect of the status characteristic was *overcome*; it was not eliminated. If race had been eliminated as a factor in expectation formation, the result

in treated groups would have been a strong inequality favoring blacks. Treatment involves giving black boys special training along with positive evaluations of performance, and then designating them "teachers" and the white boys "learners," while the blacks show the whites how to do a task. Finally, all boys are shown a video tape of the training session, and the experimenter points out instances of blacks' superior knowledge. These experiences must make a strong impression on the boys, so if it were not for the continued operation of the race status characteristic, Cohen and Roper (1972) would not have observed equality in treated groups. Of course Cohen and Roper were not concerned to produce black superiority; their goal was racial equality. But for understanding the effects of their procedure, it is crucial to know whether they overcame the effects of low status by adding several high statuses, or whether they eliminated the effects of race. Our interpretation is that they did the former.

To summarize, evidence from these three different groups of studies supports the idea that status characteristics are combined, with final expectations reflecting all available status information. In 1974 two somewhat different, explicit theories were published (Berger and Fisek, 1974; Kervin, 1974). These studies constituted the first major revision and extension of the 1966 theory to deal with multicharacteristic situations. Both theories include a combining process. In the language of our earlier example, "She may be a woman, but she's a smart woman, because she's a doctor."

The 1974 theories modified the state of knowledge concerning effects of status characteristics on social interaction as follows:

1. Rather than requiring a separate theory for each observed effect, a general and uniform explanation had been developed to account for interaction effects of such diverse characteristics as race, age, sex, and occupation. Formulation of the concept "diffuse status characteristic" recognizes that what is significant about these characteristics is that they carry with them both specific and general expectations. Formulation of specific status

status position in the group. Fisek and Ofshe (1970) found status struggles in only about half of 59 Bales groups. In the rest, some external status characteristics of members seemed to order their interaction. Such characteristics as individual's revealing a high GPA or membership in a prestigious fraternity may have led to interaction dominance, while a Southern accent, living in a dormitory, and not owning a car were likely to produce subordination. The principle seems to be: when some external status characteristic, either a D or C, differentiates members, it produces consistently differentiated expectations at the very outset of interaction, and these determine the power and prestige structure of the group without a status struggle.

¹⁰ In addition, Cohen and Roper's (1972) treatment was somewhat more successful when the raised expectations were explicitly said to be relevant to the group's task. Thus their findings parallel those of Berger et al. (1972) in this respect; please see fn. 5 above.

characteristic also is important, as its definition recognizes that there is status significance attached to performing well at tasks. Being a famous author connotes more than writing skill; it is an invidious distinction.

2. In multiple characteristic situations, the evidence then showed that all available status information, diffuse as well as specific, was combined. The elimination process often pointed to in work on race discrimination and status consistency was not supported either theoretically or empirically. However we lacked information as to the exact *form* of the combining process; no single function for describing just how status information is combined was satisfactory for all the available data.

3. The theory of status characteristics and expectation states had been tested and confirmed in a wide range of controlled and natural settings, with varied subject populations and types of interaction.

4. Along with the theory development, a most significant result of this work was developing a standardized experimental situation, the basic expectation experiment, for studying status generalization processes. A major advantage from developing the basic expectation experiment is that data from different experiments using the same subject pool usually are directly comparable.

Elimination Revisited

As the 1974 theories were in press, Freese and B.P. Cohen (1973) presented a theory for eliminating status generalization. In outline, their theory argues that it is possible to block expectation formation from a diffuse status characteristic if individuals possess at least two specific status characteristics whose evaluations differ from the evaluation of their diffuse characteristic. The reasoning behind this claim is that it is cognitively easier to infer ability from C's, which directly represent ability information, than from D's, which are less directly connected to abilities. Freese and Cohen also presented results from an experiment conducted with female college students to test their theory. In the relevant conditions 5 and 6,

subjects possessed either the high state of D and the low state of two related C's (condition 5), or the low state of D and the high state of two C's (condition 6). P(s) for condition 5 was .59, and for condition 6 it was .69, which supports an interpretation that expectations were formed only from the C's in both conditions. Thus these results differ from the 12 experiments reviewed above by showing evidence for an elimination process.

Zelditch et al. (1975) attempted to pursue this anomalous result with some experiments in which elimination ought to have a great chance to occur.¹¹ In these experiments, subjects were first told that they differed on the specific characteristic necessary to perform the group task (C*), and then given information on a differently evaluated D. Since they already had all the information needed to form task-specific expectations from C*, it should be easier to eliminate D in this situation than in any other.

Experiments used female and male college students as subjects, and produced mixed results. In the first series, combining was found for one condition (low D/high C), but elimination was found in the corresponding high D/low C condition. The latter condition was replicated with a different college student population, this time finding support for a combining process. On balance, then, the Zelditch et al. (1975) experiment produces more support for combining than for elimination, but cannot say that elimination (under certain, presently unknown conditions) never occurs.

Finally, two recent experiments (Webster, 1977) provide information on cases in which status characteristics do not function in the burden of proof process; namely, when they equate actors. For instance, noticing that oneself and another juror are both *white males* does not lead to formation of equal performance expectations for self and other. The two previous versions of the theory were indeterminate with respect to whether equating char-

¹¹ Data from these experiments, along with a somewhat different interpretation than the one given here, are published in Alexander and Lauderdale (1977).

acteristics provide structuring information. The evidence from Webster's experiments is that they do not.

The Third Theory, and a Test

In 1977, Berger et al. presented an extensive revision of the theory, incorporating the above results. Major theoretical developments are as follows:

1. It applies to situations with any number of D and C status characteristics. In multicharacteristic situations, the theory predicts aggregate expectations will be formed by combining, and this prediction now is derived from explicit assumptions about cognitive information processing.
2. It predicts to situations including some actors not in immediate interaction. Noninteractants, called "referent others," can, under specified circumstances, affect the expectations and behavior of interactants.
3. It permits predictions in sequential interaction, with actors entering and leaving the interaction situation.
4. The logical rigor and determinacy of prediction of the theory are improved over previous versions.
5. Status information which is explicitly made relevant—as it was, for example, in the Berger et al. (1972) and the Cohen and Roper (1972) experiments—is predicted to have greater effect on performance expectations than status information for which relevance must be inferred.
6. Along with the substantive theory, a graph-theoretic model is presented. This model permits prediction of numerical values of P(s) rather than ordinal predictions in the standard experimental situation, thus allowing stronger tests of the theory.

Besides incorporating several theoretical developments, this third version of the theory accords well with previous data. With the exception of conditions 5 and 6 of the Freese-Cohen experiment, results of all natural and experimental studies are

explainable. Our purpose now is to describe an experiment conducted after the theory was formulated, which tests some central aspects of it. We selected an experiment relevant to both a theoretical issue, combining vs. elimination, and an applied issue, mixed race interaction.

The experimental situation involved a single diffuse status characteristic, race, and two specific status characteristics, which were standard laboratory tests. No characteristic was made explicitly relevant to the phase 2 criterion task.¹² Subjects were white female students, paid volunteers recruited from classes at the University of South Carolina. The experiment had three conditions, differing in the amount and type of status information available to the subject. In condition 1, high D, they were shown their partner on closed circuit TV, a black female university student. In condition 2, high D/low C's, they saw their partner, were given tests purported to measure the laboratory abilities, and scores were announced. The subject scored low on both tests, and her black partner scored high. In condition 3, low C's, subjects never saw their partners but were given the same tests and scored as in condition 2. Postsession interviewing established that subjects in condition 3 assumed their partners were white.

The sequence of the experiment was as follows. Each subject entered a small laboratory room containing closed circuit TV for communication and a machine with pushbuttons and lights to indicate her own choices and to see her purported partner's choices. In phase 1 the host experimenters explained that each subject was part of a two-person group, and that the two of them would be making decisions on a variety of tasks under individual and team conditions. All status information was then introduced—D (in conditions 1 and 2) by repeatedly showing the subject a black partner on TV, and C's (in conditions 2 and 3) by administering and scoring

¹² C₁ was meaning insight, the ability to match meanings of English and primitive language words. C₂, relational insight, is the ability to match English word sounds with sounds of ancient Japanese characters. The phase 2 data collection task, deciding which of two patterns contains the greater proportion of white area, is contrast sensitivity.

laboratory tests for the subject and her partner.

For phase 2, the experimenters announced that the subject and her partner would be working on a new task, contrast sensitivity, and they would be working as a team. Each time a slide appeared on the screen, both subjects would indicate initial choices, and this information would be exchanged between them. After that, they would restudy the slide and consider their partner's initial choice, and then make their final choice for each slide. Only final choices were to be counted in determining scores for phase 2; initial choices were only for exchanging information. Experimentally controlled disagreements were introduced on 20 of 23 initial choices, and the $P(s)$ statistic—assumed to measure relative aggregate expectations—was calculated as the proportion of times each subject made the same final as initial choice.

Since the theory claims all differentiating status information will be used in forming aggregate expectations, it predicts that conditions of this experiment will be ordered as follows in terms of $P(s)$: $1 > 2 > 3$. If the diffuse characteristic race were eliminated in forming expectations when specific characteristics are present, as the Freese-Cohen theory would predict, then conditions should be ordered: $1 > 2 = 3$. If race were overlooked entirely, or if race were not a diffuse status characteristic for these subjects, conditions also would be ordered $1 > 2 = 3$. A fourth possibility, a racist prediction, is that ability characteristics will be ignored when racial differences are present. This would produce the ordering $1 = 2 >> 3$. Table 1 presents data from this experiment, and Table 2 gives Mann-Whitney U and Jonckheere tests of differences between conditions.

Results in Tables 1 and 2 conform well

Table 1. $P(s)$ Data by Conditions, Experiment on Overcoming Status Generalization

Condition	$P(s)$	σ^2	N
1	.677	2.83	22
2	.583	16.93	21
3	.508	12.45	20

Note: Variance is calculated about the mean number of stay responses, not $P(s)$.

Table 2. Statistical Tests of Predicted Ordering of Conditions

Prediction	U or (τ)	Z	p
$1 > 2$	169.5	1.51	$\approx .03$
$1 > 3$	70.0	3.82	$< .01$
$2 > 3$	157.5	1.37	$\approx .04$
$1 > 2 > 3$	(.3994)	3.3312	$< .01$

Note: U tests are used for two-condition predictions. Jonckheere (1954) test is used for the three-condition prediction, and the statistic generated is Kendall's tau.

to the first predicted ordering, that of the theory, better than to any of the alternatives. Table 1 shows that the conditions are ordered $1 > 2 > 3$, and the differences between adjacent conditions are about equal ($\approx .09$). Table 2 shows that the differences between conditions are substantial, and that the predicted ordering is sustained by the statistical test.

A more stringent test of this theory is possible using the graph model to predict exact numerical values of $P(s)$, rather than ordinal relations, for this experiment.

To predict $P(s)$ for these three conditions, we use the following formula:¹³

$$P(s) = m + q(e_p - e_o)$$

where e_p is the expectations held for the actor p ;

e_o is the expectations held for the actor o ;

$(e_p - e_o)$ is p 's expectation advantage;

m is a population parameter, representing individuals' baseline propensity to reject influence in this situation (independent of their expectations);

q is a situational parameter representing features of the experiment which increase or decrease the effects of expectations (such as greater or less task orientation, greater or less motivation to find the correct answer, etc.).

Berger et al. (1977) estimate m by averaging the $P(s)$ from two mirror image conditions of an experiment, such as one with a single high D and one with a single

¹³ In the interest of brevity, we present here only those parts of the theory used to calculate data for this experiment, in order to enable readers to check our derivations. For the complete theory, please see Berger et al., 1977.

low D. Presumably all effects of status information cancel out when this is done, leaving only an estimate of subjects' general tendency to reject influence under disagreement. But since our experiment did not have mirror image conditions, we estimate the parameters m and q by a least squares technique, substituting the observed values of $P(s)$ and the calculated values of e_p and e_o in the above formula for each condition.¹⁴ There are three linear equations with two unknowns, and this system of equations can be solved for the unknown parameters using least-squares minimization. This yields $m = .6202$, and $q = .1316$. Estimating two parameters from three data points leaves one degree of freedom for testing the model.

To calculate e_p and e_o , we need two things. First is a way to quantify the effect of the D and C characteristics upon each actor's expectations, and second is a theoretical assumption telling how these effects are aggregated for conditions 2 and 3 of our experiment. Numerical values for effects of D and C characteristics with differing degrees of relevance to the criterion task in this experiment may be calculated using parameter estimates provided in Berger et al. (1977:141-4). Calculation details for this experiment are provided here in the Appendix.

Determining the aggregate (combined) effect of two or more status characteristics requires recourse to three theoretical assumptions about how individuals process status information. These are (1) the inconsistency effect; (2) the attenuation effect; and (3) the principle of organized subsets. Briefly, the inconsistency effect argues that a single piece of status information which does not fit into the overall picture of status characteristics (such as a single negative characteristic in

a field of positive characteristics, or vice versa) has more effect on aggregate expectations than it would by itself. The attenuation principle says that each additional unit of consistent status information has less importance in determining aggregate expectations than the previous unit. For example, learning that a person already known to be a white, adult, male, and a college graduate also has a managerial job adds very little to the high expectations already formed by the race, age, sex, and education characteristics. The principle of organized subsets says that the status organizing process works in this order: first, all negative status elements are combined according to the attenuation principle; next, all positive elements are combined according to the same principle; and then the positive and negative status sets are combined. These three assumptions are not simply calculation rules; they incorporate theoretical ideas about how individuals deal with status information.

Table 3 presents the status information and calculated expectations for this experiment, and Table 4 presents the theoretically predicted $P(s)$ data, using the quantities of Table 3.

Predictions of the model are quite close to the observed values for all three conditions. Errors are all less than .01, and the predicted values deviate considerably less than one standard error from the observed values. Thus we conclude that the fit of the model to our data is satisfactory.

Additional evidence on the theory is provided by a technique developed by Zeller and Warnecke (1973), which attempts to obtain a measure of a previously unobserved step in the theoretically assumed burden of proof process. Zeller and War-

Table 3. Theoretical Quantities in the Experiment

Con- dition	Status Information	Expectations
1	p: D ₁ + o: D ₁ -	$e_p = +.2284888$ $e_o = -.2284888$
2	p: D ₁ +, C ₁ -, C ₂ - o: D ₁ -, C ₁ +, C ₂ +	$e_p = -.1762817$ $e_o = +.1762817$
3	p: C ₁ -, C ₂ - o: C ₁ +, C ₂ +	$e_p = -.4047705$ $e_o = +.4047705$

$m = .6202$ $q = .1316$

¹⁴ Variance in condition 1 (shown in Table 1) is extremely low for experiments using this basic situation. In other research, reported variance is normally in the range 12-25 exhibited by our conditions 2 and 3, with a previous low of 4.73 (Berger-Fisek, 1970) and a high of 46.24 (Berger et al., 1972). It seems likely that the variance of our condition 1 itself represents experimental variance. We thank Professor Hamit Fisek for suggesting this method of parameter estimation.

Table 4. Predicted P(s) Data, Using Point Estimation Calculations

Con- dition	Pre- dicted	s.e.	D	D /s.e.
1	.6803	.014	.0033	.2357
2	.5738	.018	.0092	.5111
3	.5136	.024	.0057	.2375

Note: s.e. is standard error of the observed P(s) values shown in Table 1. |D| is the absolute magnitude of difference between predicted and observed P(s) values.

necke point out that the theory assumes a simple three-step causal model in which:

$$D_1 \rightarrow \text{GES}_1 \rightarrow C^*;$$

or in words, a salient diffuse status characteristic (in this case, race) produces a general expectation state (a set of ideas regarding doing well or poorly at several tasks), which in turn produces task specific expectations for contrast sensitivity (measured in the experiment by P(s)). The Zeller-Warnecke technique attempts to measure the unobserved construct general expectation state through a questionnaire asking the individual to rate both herself and her partner on several dimensions of success and competence.

Using this technique in their experiment, Zeller and Warnecke were able to show a considerable increase in explained variance in their outcome variable, comparable to our P(s). We administered a modified form of the Zeller-Warnecke questionnaire to subjects after the phase 2 interaction.¹⁵

Correlation between status characteristic information and the Zeller-Warnecke measure, or between the measure and P(s), was quite high in all conditions (> .89), so it does not add appreciably to explained variance in this experiment. However it does provide an independent test of the theory, in that we predict the expectations measured in this way to be

ordered just as the P(s) measure is ordered. Responses on the questions ranged from 1 (much more success for self than partner) to 5 (much more success for partner than self). The average scale points checked in the three conditions were:

condition 1: 2.56; condition 2: 3.13; condition 3: 3.70.

The Jonckheere test shows that these differences are reliable, so the predicted ordering is sustained ($Z = 5.97$, $p < .01$).

The Zeller-Warnecke technique also gives a way to assess whether in fact an experimenter was successful in the attempted status manipulation. That is, was the characteristic chosen for manipulation actually a *status* characteristic (carrying invidious evaluations and a general expectation state) for the subject population at the time of the experiment? The ordering of experimental conditions indicates that race (D_1) and the two laboratory abilities (C_1 and C_2) were status characteristics here. It seems fair to conclude that these tests of the theory provide confirmation. All three sets of data—ordinal predictions of P(s), point estimates of P(s), and the ordering of the Zeller-Warnecke measure of GES—are as predicted by the theory.

Something these results do not tell is exactly how this outcome is to be compared to the results of Freese and Cohen (1973). In view of the pattern of other experimental studies, it does not seem likely that the Freese-Cohen results were due to any special properties of the status characteristics they used, nor to peculiarities of their subject population. It is conceivable that something about their experimental procedures blocked status generalization, but there is little information as to what might have produced that effect, since it did not occur reliably in the Zelditch et al. (1975) experiments, in the other studies reviewed, nor in our experiment reported above. But rather than dwell on that unexplained exception, we prefer to emphasize what we do know: that, in general, expectation states form by combining all differentiating status information, and that the third version of the theory did a good job of predicting the outcome of this combining process in the first independent test of it.

¹⁵ Questionnaire is available from the authors on request. Zeller and Warnecke found it necessary to modify their first measure when factor analysis showed it to contain three independent factors. We adopted their most satisfactory questions and added others; in this, we were guided by the theoretical idea of a general expectation state. Factor analysis showed our measure to be unidimensional, with average interitem correlation = .482. (Zeller-Warnecke modified measure interitem correlation = .437.)

Summary and Conclusions

We began this review by posing the classical sociological question of how individuals' statuses affect their interaction patterns, when the statuses are not known to be directly relevant to their immediate task. The first general answer to that question was provided by the 1966 theory, which formulates the burden of proof assumption, explaining how a status characteristic produces expectations which then determine interaction patterns. Experiments by Moore (1968) with college females differentiated by the status characteristic academic achievement, and by Berger et al. (1972) with Air Force sergeants differentiated by rank, support that formulation.

The second step was to extend the theory to deal with multicharacteristic situations. To do this, diffuse and specific status characteristics were distinguished, and the revised theories predicted that all available status information would be combined somehow in forming expectations (Berger and Fisek, 1974; Kervin, 1974). The majority of experiments both before and after the second theories support a combining hypothesis.

The third step (Berger et al., 1977) presents some more completely developed ideas on how individuals process status information. Our experiment, the first direct test of predictions from that theory, provides good support.

For sociologists primarily concerned with natural situations of status generalization, we now mention a few applied uses of this knowledge. First, the evidence seems clear that actors usually are not treated solely in terms of their diffuse status characteristics. That is, there is no theoretical or empirical support for a contention that "No matter how many skills or how much experience I have, they treat me as just another (woman, black, child)." Rather, what happens is better represented by a statement such as "She may be a woman, but she's a smart woman—perhaps not as smart as a man with equivalent training, but smarter than most women." Overall, there is strong evidence that both diffuse and specific characteristics form the basis for the expectations assigned to an actor.

Second, we know that to overcome the effects of a diffuse status characteristic, it is not sufficient to change expectations of only the minority (D-) individuals. To do that is more likely to create a status struggle and increase intergroup hostility than anything else. Both groups, those with the high state of D and those with the low state of D, must change their expectations in order to modify the interaction disability of minority group members.

Third, the theory suggests ways to deal with situations in which status generalization occurs on the basis of diffuse status characteristics. Under controlled conditions such as in the laboratory, an experimenter may be able to block the burden of proof simply by using his authority to tell individuals that the D is not relevant to the task at hand. Saying "We know from past research that sex (or race) is not relevant to contrast sensitivity; that is, some men (or whites) do well and some do poorly, as some women (or blacks) do well and some do poorly," may be sufficient to block expectation formation on the basis of sex or race. One natural analog to this case might be in the juries Strodtbeck and Mann (1956) and Strodtbeck et al. (1958) studied. If a judge told jurors that he knew from experience that men and women (or blacks and whites) make equally effective jurors, this might block burden of proof.

A different technique to make status characteristics irrelevant is to publish information on actual task capacities. For example, some universities publish data showing the grade point averages and the number of successful graduates among opportunity students (usually, racial minority students). In most of these reports, opportunity students do about as well as majority (high D) students at the university. Dissemination of actual performance can do much to remove the general expectations associated with race; that is, to remove race from one of the essential criteria for being a diffuse status characteristic. Similarly, corporate advertising on television showing women successfully completing difficult jobs should have the same effect for the characteristic sex.

Fourth, when the burden of proof cannot be broken for some reason, interven-

tion should take as its goal overcoming status generalization, not eliminating it, since in most situations the latter probably is impossible. To overcome status generalization, the procedure followed in our condition 2, or in the E.G. Cohen and Roper experiments, has been shown to be effective. This procedure is to add on characteristics, to demonstrate superior—not equal—performance, on several specific status characteristics by group members with the low state of D. Because the low state of the diffuse status characteristic tends to produce low expectations, producing equal expectations, and thus interaction equality, requires introducing some counterbalancing positive information. To get blacks and whites to interact as equals requires demonstrating black superiority at some specific tasks—or at least this will be true so long as race is a diffuse status characteristic in our culture.

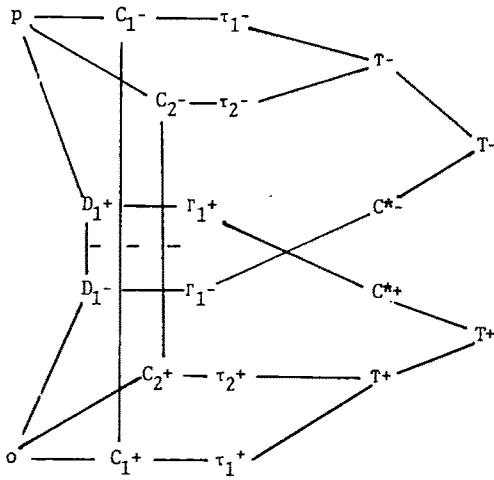
Finally, in concluding this review it seems appropriate to note that we know much more about how status elements of different types affect interaction than we did at the time of the jury studies and biracial work group experiments of two decades ago. We know about burden of proof, the interaction effects of expectations, and about multicharacteristic situations where the characteristics are either consistent or inconsistent. What we do not know at this point is exactly what conditions, if any, can completely eliminate status generalization.

REFERENCES

- Alexander, C.N. and P. Lauderdale
1977 "Situating identities and social influence." *Sociometry* 40:225-33.
- Anderson, N.H.
1968 "A simple model for information integration." Pp. 731-43 in R.P. Abelson, E. Aronson, W.J. McGuire, T.M. Newcomb, M.J. Rosenberg, and P.H. Tannenbaum (eds.), *Theories of Cognitive Consistency: A Sourcebook*. Chicago: Rand McNally.
1977 "Some problems in using analysis of variance in balance theory." *Journal of Personality and Social Psychology* 35:140-58.
- Archibald, W.P.
1976 "Face-to-face: the alienating effects of class, status, and power differences." *American Sociological Review* 41:819-37.
- Berger, J., B.P. Cohen and M. Zelditch, Jr.
1966 "Status characteristics and expectation states." Pp. 29-46 in J. Berger, M. Zelditch, Jr., and B. Anderson (eds.), *Sociological Theories in Progress*, Vol. 1. Boston: Houghton Mifflin.
1972 "Status characteristics and social interaction." *American Sociological Review* 37:241-55.
- Berger, J. and T.L. Conner
1969 "Performance expectations and behavior in small groups." *Acta Sociologica* 12:186-97.
- Berger, J., T.L. Conner and W. McKeown
1969 "Evaluations and the formation and maintenance of performance expectations." *Human Relations* 22:481-502.
- Berger, J. and M.H. Fisek
1970 "Consistent and inconsistent status characteristics and the determination of power and prestige orders." *Sociometry* 33:287-304.
1974 "A generalization of the theory of status characteristics and expectation states." Pp. 163-205 in J. Berger, T.L. Conner, and M.H. Fisek (eds.), *Expectation States Theory: A Theoretical Research Program*. Cambridge, Ma.: Winthrop.
- Berger, J., M.H. Fisek, R.Z. Norman and M. Zelditch, Jr.
1977 *Status Characteristics and Social Interaction: An Expectation States Approach*. New York: Elsevier.
- Caudill, W.
1958 *The Psychiatric Hospital as a Small Society*. Cambridge, Ma.: Harvard University Press.
- Cohen, E.G.
1972 "Interracial interaction disability." *Human Relations* 25:9-24.
- Cohen, E.G., M.E. Lockheed and M.R. Lohman
1976 "The center for interracial cooperation: a field experiment." *Sociology of Education* 49:47-58.
- Cohen, E.G., M. Lohman, K. Hall, D. Lucero and S. Roper
1970 "Expectation training I: altering the effect of a status characteristic." Technical Report #2, School of Education, Stanford University.
- Cohen, E.G. and S. Roper
1972 "Modification of interracial interaction disability: an application of status characteristic theory." *American Sociological Review* 37:643-5.
- Fisek, M.H. and R. Ofshe
1970 "The process of status evolution." *Sociometry* 33:327-46.
- Freese, L.
1970 *The Generalization of Specific Performance Expectations*. Ph.D. dissertation, Department of Sociology, Stanford University.
1974 "Conditions for status equality." *Sociometry* 37:147-88.
1976 "The generalization of specific performance expectations" *Sociometry* 39:194-200.
- Freese, L. and B.P. Cohen
1973 "Eliminating status generalization." *Sociometry* 36:177-93.

- Hornung, C.A.
1977 "Social status, status inconsistency, and psychological stress." *American Sociological Review* 42:623-38.
- Hughes, E.C.
1945 "Dilemmas and contradictions of status." *American Journal of Sociology* 50:353-9.
- Jackson, E.F.
1962 "Status consistency and symptoms of stress." *American Sociological Review* 27:469-80.
- Jackson, E.F. and P.J. Burke
1965 "Status and symptoms of stress: additive and interactive effects." *American Sociological Review* 30:556-64.
- Jonckheere, A.
1954 "A distribution-free K-sample test against ordered alternatives." *Biometrika* 41:133-45.
- Katz, I.
1970 "Experimental studies in Negro-white relationships." Pp. 71-117 in L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, Vol. 5. New York: Academic Press.
- Katz, I. and L. Benjamin
1960 "Effects of white authoritarianism in biracial work groups." *Journal of Abnormal and Social Psychology* 61:448-56.
- Katz, I. and M. Cohen
1962 "The effects of training Negroes upon cooperative problem solving in biracial teams." *Journal of Abnormal and Social Psychology*, 64:319-25.
- Katz, I., J. Goldston and L. Benjamin
1958 "Behavior and productivity in biracial work groups." *Human Relations* 11:123-41.
- Kervin, J.B.
1972 *An Information Processing Model for the Formation of Performance Expectations in Small Groups*. Ph.D. dissertation, Department of Social Relations, Johns Hopkins University.
- 1974 "Extending expectation states theory: a quantitative model." *Sociometry* 37:349-62.
- 1975 "Equating information and influence: a test of three models." Mimeo. Department of Sociology, University of Toronto.
- 1977 "An information combining model for expectation states theory: derivation and tests." *Journal of Mathematical Sociology* 5:199-214.
- Lenski, G.
1954 "Status crystallization: a non-vertical dimension of social status." *American Sociological Review* 19:405-13.
- 1956 "Social participation and status crystallization." *American Sociological Review* 21:458-64.
- 1966 *Power and Privilege*. New York: McGraw-Hill.
- Lewis, G. H.
1972 "Role differentiation." *American Sociological Review* 37:424-34.
- McCranie, E. and J.C. Kimberly
1973 "Rank inconsistency, conflicting expectations, and injustice." *Sociometry* 36:152-76.
- Meeker, B.F. and P.W. O'Neill
1977 "Sex roles and interpersonal behavior in task oriented groups." *American Sociological Review* 42:91-105.
- Moore, J.C.
1968 "Status and influence in small group interaction." *Sociometry* 31:47-63.
- Park, R.E.
1925 "The concept of position in sociology." *American Sociological Review* 20:1-14.
- Park, R.E. and E. Burgess
1921 *Introduction to the Science of Sociology*. Chicago: University of Chicago Press.
- Sampson, E.E.
1963 "Status congruence and cognitive consistency." *Sociometry* 26:146-62.
- Sherif, M., B.J. White and O.J. Harvey
1955 "Status in experimentally produced groups." *American Journal of Sociology* 60:370-9.
- Simmel, G.
[1908] *Soziologie, Untersuchungen über die Formen der Vergesellschaftung*. Pp. 87-408 in K. Wolff (trans. and ed.), *The Sociology of Georg Simmel*. New York: Free Press.
- Strodtbeck, F.L., R. James Simon and C. Hawkins
1958 "Social status in jury deliberation." Pp. 379-88 in E.E. Maccoby, T.M. Newcomb, and E.L. Hartley (eds.), *Readings in Social Psychology*. New York: Holt.
- Strodtbeck, F.L. and R.D. Mann
1956 "Sex role differentiation in jury deliberations." *Sociometry* 19:3-11.
- Torrance, E.P.
1954 "Some consequences of power differences on decision making in permanent and temporary three-man groups." Pp. 600-9 in A.P. Hare, E.F. Borgatta, and R. F. Bales (eds.), *Small Groups*. New York: Knopf.
- Webster, M.
1975 *Actions and Actors: Principles of Social Psychology*. Cambridge, Ma.: Winthrop.
- 1977 "Equating characteristics and social interaction: two experiments." *Sociometry* 40:41-50.
- Webster, M., L. Roberts and B. Sobieszek
1972 "Accepting significant others: six models." *American Journal of Sociology* 78:576-98.
- Webster, M. and B. Sobieszek
1974 *Sources of Self-Evaluation*. New York: Wiley.
- Zelditch, M., P. Lauderdale and S. Stublarec
1975 "How are inconsistencies in status and ability resolved?" Technical Report #54, Laboratory for Social Research, Stanford University.
- Zeller, R.A. and R. Warnecke
1973 "The utility of constructs as intervening variables in the interpretation of experimental results." *Sociological Methods and Research* 2:85-110.

APPENDIX
Calculating P(s) Values for
This Experiment



C₁=Meaning Insight
C₂=Relational Insight
D₁=race

paths:

p: 4+, 5+, 4-, 4-, 5-, 5-

o: 4-, 5-, 4+, 4+, 5+, 5+

Figure 1. Graph-Theoretic Representation of Condition 2

Figure 1 presents the graph for condition 2 of this experiment, which is the most complex status situation. For condition 1, all paths connecting p and o to T⁺ and T⁻ going through C₁ or C₂ would be removed. For condition 3, all paths connecting p and o to T⁺ or T⁻ and going through D₁ are removed. Both p (subject) and o (taped partner) possess positively and negatively evaluated states, respectively, of the diffuse status characteristic D₁. P and o have induced or inferred a link to a general expectation state associated with D₁, represented as T₁. The element C* is the ability needed to complete their task T; in this case, contrast sensitivity. The link between T₁ & C* represents the burden of proof process. The

negative link between D₁⁺ and D₁⁻, and the complementary states of C₁ and C₂, is dimensionality, and represents their opposite valences.

Paths through C₁ and C₂ represent the additional status information given about meaning insight and relational insight. When individuals notice C's, they induce or think about the idea of tasks for which the C's are relevant; when C₁ becomes salient, it induces the idea of working meaning insight problems, which is represented by τ₁. Note that C's are linked to τ's having the same signs. Next, working meaning insight problems induces the idea of working tasks in general. This is represented by τ, and since it applies to any type of task, it does not have a subscript. The idea of task completion induces the idea of doing the specific group task, represented by T⁺ (success) or T⁻ (failure). The graph is completed when all actors are connected to states of T using all initially given status information.

Expectations associated with an actor are a function of (1) the number of paths connecting that actor to task outcomes (T⁺ and T⁻); (2) the positive or negative signs of paths; and (3) the lengths of paths. To calculate sign, we multiply signs on each link between elements (the vertical dimensionality link is the only negative link here), and finally multiply this by the sign of the task outcome (T⁺ or T⁻). The length of a path is the number of its links between the actor and T, and number of paths is simply the total of different ways to connect each actor to T.

The attenuation principle, the inconsistency principle, and the principle of organized subsets are realized in the following formulae for calculating expectations e_p and e_o:

$$e_x = \begin{cases} e_{x+} & \text{if the paths are positive} \\ e_{x-} & \text{if the paths are negative} \end{cases}$$

where

$$e_{x+} = 1 - (1 - f(i)) \dots (1 - f(n)) \text{ and}$$

$$e_{x-} = -\{1 - (1 - f(i)) \dots (1 - f(n))\}$$

$$\text{and } e_x = (e_{x+}) + (e_{x-}).$$

Functions for paths of varying lengths, the f(i) values, are in Berger et al., 1977:144. Numbers needed here are: f(4) = .1768; f(5) = .06279. In Figure 1 of condition 2, p has positive paths of lengths 4 and 5, two negative paths of length 4, and two negative paths of length 5. In condition 1, p has two positive paths of lengths 4 and 5, and in condition 3, p has two negative paths of length 4 and two negative paths of length 5. The f(i) values are substituted in the equations above to obtain the expectations and then the predicted P(s) values of Table 4.

SEX DIFFERENCES IN CHILDREN'S FRIENDSHIPS*

DONNA EDER

University of Wisconsin, Madison

MAUREEN T. HALLINAN

Stanford University

American Sociological Review 1978, Vol. 43 (April):237-250

This paper examines sex differences in the exclusiveness of children's dyadic friendships. First, differences in the frequencies of triad types representing exclusive and nonexclusive dyadic friendships are compared. As expected, girls are found to have a greater frequency of exclusive triads than boys while boys are found to have a higher percentage of nonexclusive triads than girls. Girls are also shown to have a higher frequency of exclusive than comparable nonexclusive triads while no such pattern exists for boys. A second analysis examines movement across triad types over time; it shows that girls tend to return to an isolated dyad while boys tend to expand their dyadic friendships to include a third person. The implications of these results for differential learning of social skills and for the experience of newcomers to a group are discussed.

Several stereotypes exist regarding sex differences in children. Many of these stereotypes pertain to social interactions and suggest, in general, that girls are more social than boys. In a recent review of the literature on sex differences in children, Maccoby and Jacklin (1974) examine the empirical studies in this area. They report that there is little or no empirical support for the beliefs that girls are more motivated to receive social rewards, have a greater need for affiliation, and are more socially oriented than boys.

Recent studies suggest that boys and girls may differ in the kinds of social experiences they have rather than in the number or importance of their social interactions. One consistent finding from a number of studies of children's play is that girls tend to interact in small groups, particularly dyads, while boys tend to interact in larger groups (Waldrop and Halverson, 1973; Laosa and Brophy, 1972; Omark and Edelman, 1973; Lever, 1974). The observation of sex differences in these interactions provides a partial basis

for expecting differences in the friendship patterns of boys and girls. More specifically one would expect girls' dyadic friendships to be more exclusive than those of boys, where exclusive dyadic friendships are defined as those in which dyad members fail to include third persons as friends.¹

While many studies (Parsons, 1966; Gronlund, 1959; Campbell, 1964) have shown that a strong sex cleavage exists during the elementary school years, no study as yet has analyzed sex differences in the friendship patterns of children. The present paper will examine sex differences in the exclusiveness of children's dyadic friendships² by comparing differences in the choice or nonchoice by dyad members of a third person at a given point in time as well as change in these triadic choices over time.

Interaction and Exclusiveness of Dyadic Friendships

Several theorists suggest that interaction is a strong predictor of positive attraction. Homans (1950) argues that the de-

* The authors' names are listed alphabetically; each shares equal responsibility for the work. The research reported here was supported in part by a National Science Foundation Graduate Fellowship and by Spencer Foundation Grant CAX 677. Computer time was provided by the University of Wisconsin, Madison. We are indebted to Diane Felmlee for special assistance with the computer programs. We are also grateful to Diane Felmlee, Robert Kaufman, Aage Sørensen, William Sewell, and two anonymous referees for helpful comments.

¹ This definition of the word "exclusive" should be distinguished from its definition as an unwillingness on the part of *individuals* to make friends. Unlike most previous sociometric studies, this paper focuses on the dyad as the unit of analysis rather than on the individual.

² The term friendship will be used throughout the paper to refer to a best friend relationship as compared with a less intense friend relationship.

gree of liking between two persons increases with frequency of interaction. According to exchange theorists (Thibaut and Kelley, 1959; Blau, 1964), interaction is ordinarily a rewarding experience with more frequent interaction bringing greater rewards. The likelihood of continued positive reinforcement leads to friendship. Heider (1958) explains the effect of interaction on friendship in terms of a structural tendency toward balanced sentiment relations in face-to-face groups. In accord with this close relationship between interaction and friendship, one would expect the friendships of children to form to a great extent through their interactions with other children.

Two explanations can be offered for the fact that boys' and girls' social interactions occur mainly in different sized groups, which in turn suggest that the dyadic friendships of boys and girls should differ in degree of exclusiveness. The first is based on differences in the number of participants in boys' and girls' games. In a study where children between the ages of nine and fifteen were asked to indicate which games they preferred, Sutton-Smith and Rosenberg (1971) found that the games selected by boys tended to require a larger number of participants than those selected by girls. The fact that boys and girls tend to prefer different play activities could be due in part to parental, teacher, and peer group reinforcement to engage in appropriate sex-linked activities (Mischel, 1970). This difference in game preference could also be explained by a tendency for children to value those things that are consistent with their self-concept of being male or female, and thus to prefer those games which they view as being typically male or female activities (Kohlberg, 1966).

If girls and boys are socialized to prefer and participate in different types of play activities which require different numbers of participants, it is likely that this will have implications for their friendship patterns. While the requirement of a large number of participants for male-type games encourages male dyads to expand, there is no similar pressure on female dyads to include third persons as friends.

Secondly, girls tend to engage in more

intimate behavior than boys and dyads are conducive to intimate behavior. According to Simmel (1950), intimate conversations are more likely to occur in dyads than in larger groups. Since one of the costs to be considered when deciding whether or not to disclose intimate information is the informational power given to the listener (Altman and Taylor, 1973), it would be less threatening to discuss intimate topics in a dyad as only one person would obtain informational power.

Two studies that examined self-disclosure among adolescents (Rivenbark, 1971; Dimond and Munz, 1967) found that girls reported disclosing more to their same-sex best friend than did boys. The tendency for girls to be more intimate with same-sex friends than boys may be due, in part, to differential reinforcement of intimate behavior. Whereas hand holding and other signs of affection are readily accepted between girls, such signs of intimacy are viewed as unmasculine and are generally discouraged between boys. Since intimate behavior is rewarding for girls, it is likely that they will want to preserve the opportunity for continued intimate conversation by excluding others from their dyadic friendships. However, since boys are not reinforced for intimate behavior, they would have less reason for insuring the opportunity for interaction in a dyad, and thus have less preference for exclusive dyadic friendships.

Both explanations for sex differences in the size of children's social groups lead us to hypothesize that the dyadic friendships of girls are more exclusive than the dyadic friendships of boys. An experiment by Feshback and Sones (1971) offers some support for this hypothesis. In a study of twelve- and thirteen-year-old children, they compared the reaction of established pairs of same-sex close friends to a newcomer. The results showed that boy pairs tended to speak to the newcomer sooner, were more likely to incorporate the newcomer's ideas, and made more positive ratings of the newcomer. The present study will test the proposition that girls' dyadic friendships are more exclusive than boys' by examining sex differences in patterns of sociometric choice.

Patterns of Friendship Choices

In order to analyze the exclusiveness of dyads it will be necessary to view the dyad in relation to a third person. How the dyad responds to that person will give some indication of how exclusive the dyadic friendship is. Thus triadic configurations will be used to compare the exclusiveness of female and male dyads.

Triads may be classified according to the number of mutual (*i* chooses *j* and *j* chooses *i*), asymmetric (*i* chooses *j* but *j* does not choose *i*) and null (neither *i* nor *j* choose each other) relations that they contain, as well as by the direction of the asymmetric choices. For example, the 102 triad contains one mutual, no asymmetric and two null relationships. This classification results in the sixteen triad types presented in Figure 1.

The first step of the analysis will be to compare the frequencies of those triads containing at least one mutual relationship. Three of these triads involve no recognition of the third person by either dyad member (the 102, 111D, and 120D

triads in Figure 2) and can thus be viewed as representing exclusive dyadic friendships. The other six triads which contain a mutual relationship involve recognition of the third person by at least one dyad member (the 111U, 201, 120C, 120U, 210, and 300 triads in Figure 3). These triads will be considered as representing nonexclusive dyadic friendships. In line with the expectation that dyadic friendships of girls will be more exclusive than those of boys, one would expect that the three triad types representing exclusive dyadic friendships would be more common for girls than boys and that the six triad types representing nonexclusive dyadic friendships would be more common for boys than for girls.

Furthermore, one would expect that for those pairs of triads having similar configurations except for the direction of choices (111D and 111U, 120D and 120U), the more exclusive triad in the pair would be more common for girls than the less exclusive triad. This would not necessarily be true for boys. Thus one would predict a greater frequency of 111D than

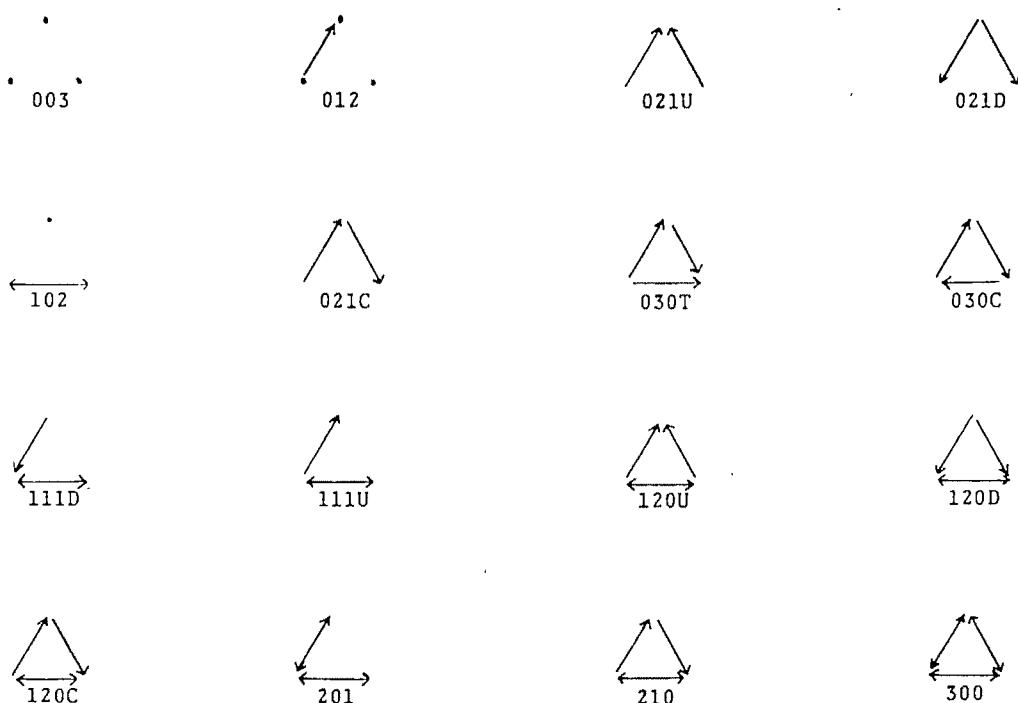


Figure 1. Sixteen Triadic Configurations Classified According to Number of Mutual, Asymmetric and Null Choices

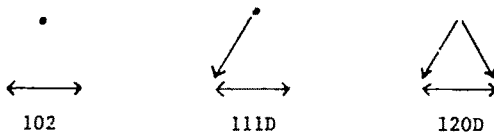


Figure 2. Triads Containing Exclusive Dyadic Friendships

111U triads and a greater frequency of 120D than 120U triads for girls than for boys.

The second stage of the analysis will involve looking at changes in triadic structures over time. The crucial triad types for this analysis are the 111D and 111U triads. Sørensen and Hallinan (1976) found these triads to play an important role in the development of group structure. They are also significant in the present study because they represent the point at which a mutual dyad first has contact with a third person and therefore provide an ideal starting point from which to analyze the response of a dyad to a third person.

Basically, there are two ways in which a dyad can respond to this initial contact from an outsider. In the case of the 111D triad, either the dyad member who was chosen can ignore the third person by failing to reciprocate his offer of friendship, or one or both dyad members can respond by choosing the third person as a friend. If the dyadic friendship is exclusive, it is likely that its members will respond in the first manner; if it is not exclusive, they should respond in the second. If the dyad member fails to reciprocate the friendship offer of the third person, the offer is likely to be withdrawn. The member then will move from a 111D triad to a 102 triad. If the dyad members respond by making further choices, the triad would change from a 111D triad to a 201 or 120C triad, depending on which person chose the outsider, or to a 210 triad if both dyad members chose the third person (see Fig. 4).

In the case of the 111U triad where a member of a mutual dyad chooses a third person, two responses are also possible.

The dyad member can withdraw his choice, perhaps because the other dyad member becomes jealous, and a 102 triad will result. Or, the second triad member, choosing the third person, can also form a 120U or a 210 triad, if the choice occurs in conjunction with a response from the third person. Movement to 102 is the expected pattern if the dyad is exclusive, while movement to 120U or 210 is predicted if the dyad is nonexclusive (see Fig. 5).³

Since it is expected that triads containing exclusive dyadic friendships will move from 111D and 111U triad types to the 102 triad type, it is hypothesized that this pattern of movement will be more common for girls than boys. Likewise, it is expected that a pattern of movement from 111D to 201, 120C and 210 and from 111U to 120U and 210 will be more common for boys.

Sociometric Data

Longitudinal sociometric data from children in one fifth-grade class (Class D) and four sixth-grade classes (A, B, C and E) will be used to test the proposition that the dyadic friendships of girls are more exclusive than those of boys. The age range in the classes was nine through twelve and the classes varied in size from 25 to 35 students. Classroom E was designated as an open classroom while the other four were more traditional teacher-centered classrooms. The classes were located in five all-white public and private schools similar in socioeconomic and ability composition.

The data were collected at seven time points throughout the school year at six week intervals beginning the second week

³ Since the 120C and 120U triads contain two asymmetric choices, they are expected to occur infrequently because sentiment relationships have a tendency toward reciprocity (Gouldner, 1960; Backman and Secord, 1959).

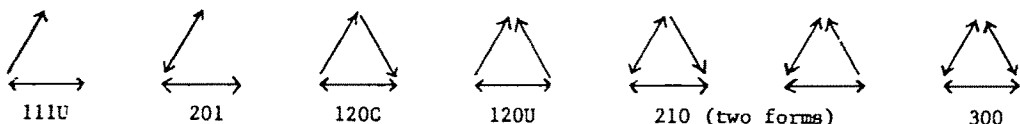


Figure 3. Triads Containing Nonexclusive Dyadic Friendships

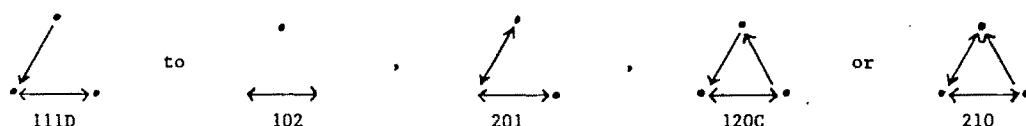


Figure 4. Some Possible Responses of Members of 111D Triad to the Friendship Choice from the Third Person to a Dyad Member

of school. The children were asked to name their best friends (the students they liked very much), and their friends (students they liked but did not consider best friends). They were allowed to choose as many or as few best friends and friends as they wished, a technique designed to minimize measurement error (Holland and Leinhardt, 1973; Hallinan, 1974). Since our interest is in the exclusiveness of best friend relationships, the best friend level is used in the analysis.⁴

Results

Observed differences in the dyadic friendship patterns of boys and girls could be due to sex differences in the number of friendship choices made, which in turn could have implications for the proportion of various triad types. Therefore, we calculated the average number of choices made by sex. Table 1 presents these data for each class across the seven time points. A t-test showed no difference in the average number of choices made by boys and girls in Classes A, B and E. In Class D, boys made significantly more choices than did girls while in Class C the opposite pattern held. The latter results are likely due to the ratio of boys to girls in these classes. The number of choices

made seems to increase with the number of persons in the group, as might be expected, because there are more children to choose from. Since the relationship between the size and number of choices given appears to be the same for both sexes, the results show no evidence of a tendency for one sex to make more friendship choices than the other.

Table 2 presents the average frequency of the nine triad types for the five classes by sex. The percentages were obtained by dividing the number of each triad type observed at a given time by the total number of same-sex triads in the class at that time and averaging the results over all the time points. Table 2 also gives the statistic $(O - E)/SD$ where O is the observed number of triads of a given type, E is the number expected by chance and SD is the standard deviation. The random model on which this statistic is based controls for the size of the group and the number of mutual, asymmetric and null relations in the group. Since these constraints impose severe restrictions on the placement of choices in a sociomatrix (Hallinan and Sørensen, 1973), this particular random model provides a conservative test of the hypothesis. Nevertheless, it is the model most frequently used to test for randomness (see Holland and Leinhardt, 1970; Leinhardt, 1972; Hallinan, 1974; Hallinan and Felmlee, 1975). A negative value of this statistic indicates that the observed sociogram contains fewer of that particular triad type than expected by chance. A positive value implies more than predicted by chance. For example, the average observed number of 102 triads among boys is 0.7, which means that the



Figure 5. Some Possible Responses of Members of 111U Triad to the Friendship Choice from a Dyad Member to the Third Person

⁴ To analyze the data we looked at choices from girls to girls and from boys to boys, excluding cross-sex choices. The existence of a strong sex cleavage in our data at the best friend level implies that analysis of within-sex choices included the great majority of all choices given. Within-sex choices involved the following percentages of total choices: 99.5% in A, 75.3% in B, 86.0% in C, 98.9% in D and 95.8% in E.

Table 1. Number of Choices Made Averaged over Seven Time Points in Five Classrooms by Sex

Class		Mean Number of Choices Made	Standard Deviation
A	Boys (n=10)	6.44	2.11
	Girls (n=20)	6.35	2.78
B	Boys (n=12)	5.60	2.57
	Girls (n=14)	4.50	2.20
C	Boys (n=10)	4.26*	3.85
	Girls (n=18)	7.38	2.99
D	Boys (n=23)	8.46*	6.56
	Girls (n=12)	3.33	2.26
E	Boys (n=13)	4.87	2.25
	Girls (n=14)	5.39	2.27

* The difference between the means is significant at .05 level by a t-test.

number of 102 triads is 7/10 of a standard deviation greater than the number expected by chance. Since the underlying distribution of each triad type is unknown, although it is likely to be normal, the statistic should not be used for comparative purposes. Its utility here is merely to show that the observed results are unlikely to have occurred by chance.

The results in Table 2 reveal that girls have a greater average frequency of 102 triads than boys, as predicted, in four of the five classes (Classes A, B, D and E). In two of the four classes (Classes A and D), the average frequency is over twice as great for girls as for boys. Since the ratio of girls to boys varies across the classes, the greater frequency of 102 triads for girls is unlikely to be due to a difference in the relative number of girls or boys in the class.⁵ Girls have a greater frequency of 111D triads than boys in all five of the classes. The percentage of 120D triads is considerably greater for girls than for boys

⁵ Ideally one should control for size when comparing percentage of triad types because the number of triads increases curvilinearly with the size of the group; thus the probabilities of the occurrence of particular triad types vary. In the present study, data limitations precluded a rigorous control for size. However, the fact that a similar pattern emerges regardless of the relative number of girls and boys in a class and that no systematic effects of size are evident suggest that the differences are due to sex rather than an artifact of group size.

Table 2. Percentage of Triad Types and Measure of Randomness Averaged over Seven Time Points for Five Classes by Sex

		102		111D		120D		111U		120U		120C		201		210		300	
Class		%	MR*	%	MR*	%	MR*	%	MR*	%	MR*	%	MR*	%	MR*	%	MR*	%	MR*
A	Boys (n=10)	2.3	(0.7)	3.5	(-0.7)	5.1	(-0.2)	6.3	(0.5)	12.3	(3.0)	5.1	(-2.2)	3.9	(-1.2)	26.1	(-1.1)	20.8	(1.8)
	Girls (n=20)	27.9	(7.8)	4.3	(-5.3)	2.8	(6.4)	6.1	(-3.0)	2.0	(3.6)	0.8	(-2.8)	1.6	(-9.1)	3.8	(1.1)	3.1	(7.7)
B	Boys (n=12)	11.7	(1.3)	6.0	(-2.4)	5.1	(1.9)	7.0	(-1.9)	4.0	(1.1)	3.4	(-1.7)	3.9	(-4.3)	14.6	(0.9)	11.4	(5.7)
	Girls (n=14)	15.9	(0.9)	9.3	(0.5)	4.2	(4.7)	4.8	(-2.4)	4.1	(4.2)	0.5	(-2.7)	2.3	(-3.1)	3.2	(0.1)	2.1	(4.2)
C	Boys (n=10)	19.5	(0.8)	7.6	(-1.1)	2.6	(1.2)	6.3	(-1.5)	2.7	(1.1)	3.1	(-0.6)	6.2	(-3.6)	12.4	(1.6)	9.1	(3.4)
	Girls (n=18)	13.3	(2.2)	9.7	(-0.8)	7.2	(9.9)	6.1	(-4.2)	2.9	(1.2)	1.6	(-4.4)	5.7	(-2.9)	7.3	(0.2)	3.7	(4.5)
D	Boys (n=23)	8.4	(-5.4)	4.1	(-7.4)	1.9	(0.7)	10.3	(3.2)	5.6	(12.4)	1.3	(-5.2)	3.1	(-4.9)	6.8	(4.9)	3.3	(11.4)
	Girls (n=12)	19.9	(1.8)	6.3	(-0.8)	4.2	(5.2)	3.9	(-2.2)	0.4	(-0.7)	0.6	(-1.3)	2.2	(-1.9)	2.6	(0.6)	1.5	(2.8)
E	Boys (n=13)	23.8	(2.3)	6.9	(-1.8)	4.9	(5.1)	7.3	(-1.0)	1.9	(0.0)	1.1	(-1.2)	4.2	(-3.7)	7.4	(1.3)	5.7	(4.7)
	Girls (n=14)	27.7	(3.3)	8.6	(0.0)	3.9	(5.1)	5.8	(-1.8)	2.0	(1.0)	1.2	(-1.5)	4.2	(-5.3)	5.9	(-0.2)	7.6	(6.5)

* Measure of Randomness.

in Classes C and D, although in the remaining three classes this triad occurs more frequently among boys. In general, these findings provide strong support for the hypothesis.

Since the remaining six triad types represent exclusive dyadic relationships, they are expected to appear more frequently among boys than among girls. The hypothesis is supported by the 111U and the 210 triads in all the classes. It is also supported by the 120C, 201 and 300 triads in Classes A, B, C and D while the differences for the 120C and 201 triads in Class E are negligible. We note in the discussion that E is an open classroom. Since group interaction is encouraged in open classrooms, girls there may tend to form exclusive relationships less frequently. Least support for the hypothesis is given by the 120U triad which occurs more frequently for boys than for girls in classes A and D only. However, the difference in the percentages for boys and girls in the remaining three classes is negligible. Overall, the six nonexclusive triads consistently occur more frequently among boys than among girls with the percentage being several times as great among boys for some of the triads in particular classes. The results give additional strong evidence that boys' dyadic friendships are less exclusive than girls.

The results shown in Table 2 also indicate that, as predicted, in four of the five classes (B, C, D, and E), girls had a higher percentage than boys of 111D than 111U triads. This tendency was not found among boys who had a higher percentage of 111D than 111U in four of the classes. Similarly, in all five classes girls had a greater percentage than boys of 120D than 120U triads. There was no pattern in the preference of boys between these two triads (three classes had a greater percentage of 120U triads and two classes had a greater percentage of 120D triads).

In comparing the magnitude of the differences between the percentage of triads of a given type for boys and girls in Table 2, we observe that these differences are greatest for the 210 and 300 triads. Since these are the two densest triads, we can examine sex differences in their frequency over time to determine whether boys are

initially involved in more of these triads than girls or whether the friendship networks of boys become denser as the year progresses. Table 3 shows the percentage of 210 and 300 triads at each of the seven time points for the five classes by sex. Sex differences in temporal patterns can be seen for both triads. The percentage of 210 triads for boys is greater than for girls in only one of the five classes at the first time point. However, the frequency of this triad type increases substantially for boys over the year in four of the five classes (Classes A, B, C and D), while the percentage for girls, remaining consistently low, increases no more than three percentage points from time one to time seven in any class. A similar pattern is found for the 300 triad. Boys have more 300 triads than girls at time one in every class and the percentages increase dramatically over the school year, especially in Classes A, B and C. This finding further suggests that boys' dyadic friendships expand to include others over time while girls continue to relate primarily in dyads.

The results of the second stage of analysis, an examination of changes in triadic configurations over time, are shown in Tables 4 and 5. Table 4 presents the conditional probabilities of movement to the 102, 120C, 201 and 210 triads given exit from the 111D triad. Table 5 shows the conditional probabilities of entrance into the 102, 120U and 210 triads given exit from the 111U triad.⁶ (The probability of moving from the 111D or 111U triad to each of the other triad types is given in Appendix A and B; here we are concerned only with the triads relevant to the argument.) Movement from a 111D or 111U triad to a 102 triad involves the withdrawal of contact between a mutual dyad and an outsider, while movement to one of the expanded triads (120C, 120U, 201 and 210) represents the inclusion of the outsider as a best friend. Table 4 shows that girls are more likely than boys to move from 111D

⁶ The data were examined to insure that the mutual choice in the 102 triad involved the same two people as in the 111D triad. This was found to be the case for 91% of the girls' triads and 95% of the boys'. Similarly, the same mutual dyad was involved in the move from 111U to 102 in 91% of the girls' triads and 97% of the boys' triads.

Table 3. Percentage of 210 and 300 Triads in Five Elementary School Classes at Seven Time Points by Sex

		T ₁	T ₂	T ₃	T ₄	T ₅	T ₆	T ₇
Class		210 Triads						
		%	%	%	%	%	%	%
A	Boys (n=10)	16.7	13.3	30.0	33.3	28.3	26.7	34.2
	Girls (n=20)	3.3	2.7	3.9	3.4	4.0	5.1	4.0
B	Boys (n=12)	10.0	13.6	8.6	15.9	20.5	14.5	19.1
	Girls (n=14)	4.1	3.3	3.0	1.4	2.7	3.0	4.7
C	Boys (n=10)	5.8	9.2	17.5	15.8	12.5	14.2	11.7
	Girls (n=18)	3.1	9.2	8.5	10.2	6.9	7.1	6.1
D	Boys (n=23)	1.7	2.4	5.2	8.5	8.9	9.1	10.2
	Girls (n=12)	1.4	3.6	2.7	2.3	0.9	3.2	4.1
E	Boys (n=13)	7.3	8.0	11.2	8.4	6.6	4.9	5.2
	Girls (n=14)	3.6	3.8	12.4	5.5	5.5	4.7	6.3
		300 Triads						
		%	%	%	%	%	%	%
A	Boys (n=10)	5.8	1.7	13.3	16.7	32.5	32.5	43.3
	Girls (n=20)	3.1	3.8	2.7	4.1	2.2	2.4	3.7
B	Boys (n=12)	6.4	5.0	5.5	18.6	13.2	20.5	10.9
	Girls (n=14)	1.1	3.8	2.5	1.4	1.6	2.5	1.6
C	Boys (n=10)	4.2	6.7	11.7	14.2	12.5	9.2	5.0
	Girls (n=18)	0.7	3.8	5.9	3.3	4.2	4.0	3.7
D	Boys (n=23)	1.2	1.0	2.3	4.8	3.8	4.5	5.8
	Girls (n=12)	0.5	0.9	1.8	3.2	0.4	2.3	1.3
E	Boys (n=13)	4.2	5.6	5.9	8.7	6.6	3.8	4.9
	Girls (n=14)	6.0	5.5	10.4	9.3	8.0	6.3	7.4

to 102 in all five classes, while Table 5 reveals that girls have a higher probability of moving from 111U to 102 in three of the five classes (Classes A, B and D). In other words, after contact with a third person, girls are more likely than boys to return to

an isolated mutual dyad. This occurs both when the outsider initiates the friendship and when a member of the exclusive dyad chooses a third person.

We next consider the conditional probability of movement from the 111D and

Table 4. Conditional Probabilities of Movement from 111D Triad to Select Other Triad Types for Five Classes by Sex

Class		102	120C	201	210	201 & 120C & 210
A	Boys (n=10)	.000	.036	.036	.429	.501
	Girls (n=20)	.397	.008	.058	.029	.095
B	Boys (n=12)	.118	.029*	.103	.132	.264
	Girls (n=14)	.252	.014	.068	.020	.102
C	Boys (n=10)	.139	.111	.083	.083	.277
	Girls (n=18)	.176	.020	.153	.052	.225
D	Boys (n=23)	.109	.062	.145	.138	.348
	Girls (n=12)	.159	.016	.190	.063	.269
E	Boys (n=13)	.302	.009	.151	.057	.217
	Girls (n=14)	.370	.021*	.144	.062	.221

* These probabilities are high due to the inclusion of a mutual dyad other than the original one (see fn. 6).

Table 5. Conditional Probabilities of Movement from 111U Triad to Select Other Triad Types for Five Classes by Sex

Class		102	120U	210	120U & 210
A	Boys (n=10)	.021	.064	.255	.319
	Girls (n=20)	.420	.028	.057	.085
B	Boys (n=12)	.029	.058	.232	.290
	Girls (n=14)	.143	.036	.048	.084
C	Boys (n=10)	.167	.000	.260	.200
	Girls (n=18)	.140	.044	.114	.158
D	Boys (n=23)	.118	.008	.130	.138
	Girls (n=12)	.178	.067	.044	.111
E	Boys (n=13)	.310	.070	.080	.150
	Girls (n=14)	.305	.000	.048	.048

111U triads to one of the expanded triads.⁷ The last column in Table 4 shows that in every class except E, boys are more likely than girls to move from 111D into one of these more inclusive friendship structures. Hence boys, more frequently than girls, respond positively to a friendship offer from a third person. Examining movement from 111D to each of the three expanded triads in Table 4 reveals the same pattern in most cases.⁸ However, these data show an unexpected sex difference. In three of the five classes (A, C and D), girls have a higher probability than boys of moving into a 201 triad; boys are more likely than girls to move into a 210 triad in all classes except Class E. (The movement into 120C is infrequent for both boys and girls, as expected, since the arrangement of the two asymmetric rela-

tionships in the triad is believed to be psychologically uncomfortable.) These results are significant because movement from 111D to 210 requires two additional choices, whereas movement to 201 demands only one. Consequently, the data indicate that among boys, both members of the mutual best friend dyad are responding to the friendship choices of the outsider; among girls, a response, when there is one, is usually on the part of only one member of the dyad.

Comparing the probability of movement from 111U to either a 120U or a 210 triad, data presented in the last column of Table 5 show that boys are more likely than girls to make further choices to a third person in all five classes. When movement to each triad is examined separately, boys are seen to have a greater probability than girls of moving from 111U to 120U in three classes (Classes A, B and E) and from 111U to 210 in all five classes. These results represent the same pattern as found in movement from 111D; that is, there is a tendency for boys' dyadic friendships to be expansive while girls' dyadic friendships are more exclusive.

The data in Tables 4 and 5 show that it is more likely for boys than girls to enter the 210 triad given exit from the 111 triads; while in some classes girls are more likely than boys to enter the 201 triad and the 120U triad. It could be argued that this finding indicates only that girls' dyadic friendships expand more slowly than boys. Girls may first move into 201 and 120U triads and then into 210 triads, whereas boys may form 210 triads im-

⁷ The same mutual dyad in the 111D triad remained in the 201 triad for 95% of both boys' and girls' triads, in the 210 triad for 91% of the girls' triads and 92% of boys' triads and in the 120C triad for 60% of the girls' triads and 83% of the boys' triads. The probabilities of movement into 120C triads, which are high due to inclusion of a mutual dyad other than the original one, are indicated in the table by asterisks.

⁸ Since there is one 210 triad that does not require a response from the dyad of a 111U triad, differential movement into that triad versus other 210 triads could affect the results. An examination of the data revealed that, overall, boys used that triad as infrequently as girls (31% of the moves). In three of the classes, the relative use of that triad was similar for girls and boys. In Class A, boys used it more frequently than did girls but not enough to affect the basic pattern of boys moving from 111U to 210 more frequently than girls. In Class C, girls used it more frequently than boys.

Table 6. Conditional Probability of Movement from 201 and 120U Triads to 210 Triad for Five Classes by Sex

Class		201 to 210	120U to 210
A	Boys (n=10)	.613	.242
	Girls (n=20)	.109	.172
B	Boys (n=12)	.327	.200
	Girls (n=14)	.026	.105
C	Boys (n=10)	.364	.357
	Girls (n=18)	.223	.216
D	Boys (n=23)	.347	.288
	Girls (n=12)	.077	.167
E	Boys (n=13)	.190	.158
	Girls (n=14)	.236	.125

mediately after leaving 111. If this were the case, girls could be involved in as many dense triadic configurations as boys at some future point in time. This possibility may be examined by comparing sex differences in movement from the 201 and 120U triads to the 210 triad. That is, assuming that girls are slower in forming 201 and 120U triads, we compare sex differences in movement into the denser 210 triad, once the 201 and 120U triads have been reached. Table 6 gives these results. In every class except E, boys have a higher probability than girls of moving from 201 to 210; similarly, boys are more likely than girls to move from 120U to 210 in every class. These results counter the argument that girls and boys are involved in the same process of forming extended friendships but at different rates and support our former conclusion

that girls' dyadic friendships remain exclusive over time.

A final analysis provides further evidence to support this conclusion. Table 7 shows the triads which are most likely to occur for boys and girls, given exit from the 201 and 120U triads. In three classes (A, B and D), when girls leave the 201 triad, they are most likely to enter a triad which contains fewer friendship choices—namely, the 111 triad. The same pattern is found with greater consistency for exit from 120U; in every class the triad most likely to form among girls has fewer choices (the 111D or 102 triad). The results for boys are exactly the opposite. In moving from 201, boys are most likely to increase choices and form 210 triads in every class; similarly, in exiting from 120U, boys most frequently form expanded triads (the 210 or 300 triad) in every class except E. These findings show that girls do not form dense friendship triads over time but rather return to exclusive mutual dyads after contact with a third person. The dyadic friendships of boys, on the other hand, become more expansive over the school year. Tables 3 to 7 present strong evidence that girls' dyadic friendships are more exclusive than boys and that girls tend to resist intrusion on a mutual best friendship over time, while boys quickly expand a mutual best friend dyad to include a third person.

Discussion

The results of comparing the frequency of triads containing mutual best friend

Table 7. Triads with Highest Probability of Formation, Given Exit from 201 and 120U Triads for Five Classes by Sex

Class		Exit from 201 to			Exit from 120U to		
		111D	111U	210	102	111D	210 300
A	Boys (n=10)			0			0
	Girls (n=20)	X			X		
B	Boys (n=12)			0			0
	Girls (n=14)		X			X	
C	Boys (n=10)			0			0
	Girls (n=18)			X		X	
D	Boys (n=23)			0			0
	Girls (n=12)	X				X	
E	Boys (n=13)			0	0		
	Girls (n=14)			X	X		

dyads at a given point in time, as well as the results of comparing patterns of triadic movement over time, support our hypothesis that girls tend to have more exclusive dyadic friendships than boys. However, it is interesting to note that the expected patterns were least strong in class E, the open classroom. One explanation for the lack of clear sex differences in class E is the fact that open classrooms, structured so as to encourage more interaction within the classroom, provide students with another major opportunity for interaction besides play activities. Since classroom activities are less likely to have the differential size requirements that play activities have, there would be less difference in the need for boys and girls to have nonexclusive friendships.

Another explanation for the lack of clear sex differences in class E is the possibility that teachers in open classrooms, as well as the parents of children who attend open classrooms, may engage in socialization practices different from those of teachers and parents of students in more traditional schools. Although self-selection has not been studied in regard to selection in open classrooms, it has been shown to be an important factor in determining which students attend different colleges and universities (Alwin, 1974; Jencks and Riesman, 1968).

The fact that the open classrooms did show a different pattern than the traditional classrooms also suggests that sex differences in friendship patterns are due mainly to socialization differences as opposed to biological differences. Since sex-role socialization has been a subject of much recent debate, it is likely that these socialization practices may change considerably in the future. Sutton-Smith and Rosenberg (1971), for example, found that the sex differences in their study of game preference were weaker than they were thirty years ago, with girls showing an increasingly greater preference for the games preferred most by boys. Thus it is important to realize that the findings reported in this paper are subject to change along with changes in socialization.

However, given that there presently appear to be clear differences in the de-

gree of exclusiveness of boys' and girls' dyadic friendships, it is important to consider the implications of these differences. One major implication is the development of different types of social skills by boys and girls. Although differences in the interaction and friendship patterns of children are initially due to socialization, it is likely that once the friendships are formed, they promote further differential interaction patterns, leading in turn to the development of different kinds of social skills. Thus, while game requirements may initially encourage boys to develop nonexclusive friendships, once these friendships are formed they will further the tendency of boys to interact in large groups. This emphasis on one kind of social interaction will lead in turn to the development of particular social skills relating to group decision making, group leadership, and other group processes. On the other hand, the development of exclusive friendships in girls furthers their tendency to interact in dyads and leads to the development of other types of social skills such as the ability to engage in intimate self-disclosure.

Another implication of the differential nature of children's friendships is more directly related to the experiences of children, especially newcomers, in a group. If girls' dyadic friendships tend to be more exclusive, it is likely that female newcomers would have greater difficulty than male newcomers making close friends. This might have a negative effect on the child since she may attribute her inability to make close friends to personal inadequacy, and not realize that female friendships tend to be more exclusive in general.

More refined analyses of children's friendships are needed to obtain a better understanding of the process of friendship formation among boys and girls. A direct comparison of the amount of intimate disclosure to same-sex friends by boys and girls would provide evidence of a relationship between disclosure and friendship exclusiveness. Similarly, an examination of the frequency and content of children's interactions during play would demonstrate the structural constraints children's games actually place on their friendships.

These empirical findings would strengthen the explanations of sex differences in friendships provided in this paper.

In summary, two rationales for sex differences in the interaction patterns of children suggested certain sex differences in their friendship patterns. Specifically, it was argued that girls' same-sex dyadic friendships would tend to be more exclusive than the same-sex dyadic friendships of boys. This hypothesis was examined by comparing the frequency of certain exclusive and nonexclusive triad types and by comparing the patterns of triadic movement averaged over six time intervals. The results supported the hypothesis and showed, in addition, that girls' dyadic friendships remain consistently exclusive over the school year while the dyadic friendships of boys expanded to include newcomers. These findings have important implications for the development of different types of social skills in children and for the differential experience of male or female newcomers to a class.

REFERENCES

- Altman, Irwin and Dalmaz Taylor
1973 *Social Penetration: The Development of Interpersonal Relationships*. New York: Holt.
- Alwin, Duane
1974 "College effects of educational and occupational attainments." *American Sociological Review* 39:210-23.
- Backman, C. W. and P. F. Secord
1959 "The effect of perceived liking on interpersonal attraction." *Human Relations* 12: 379-84.
- Blau, Peter
1964 *Exchange and Power in Social Life*. New York: Wiley.
- Campbell, John
1964 "Peer relations in childhood." Pp. 289-322 in M. L. Hoffman and L. W. Hoffman (eds.), *Review of Child Development Research*, Vol. 1. New York: Sage.
- Dimond, R. E. and D. C. Munz
1967 "Ordinal position of birth and self-disclosure in high school students." *Psychological Reports* 21:829-33.
- Feshback, N. D. and G. Sones
1971 "Sex differences in adolescent reactions toward newcomers." *Developmental Psychology* 4:381-6.
- Gronlund, Norman
1959 *Sociometry in the Classroom*. New York: Harper Brothers.
- Gouldner, A. W.
1960 "The norm of reciprocity: a preliminary statement." *American Sociological Review* 25:161-78.
- Hallinan, Maureen
1974 *The Structure of Positive Sentiment*. New York: American Elsevier.
- Hallinan, Maureen and Diane Felmlee
1975 "An analysis of intransitivity in sociometric data." *Sociometry* 38:195-212.
- Hallinan, Maureen and Aage Sørensen
1973 "Measuring transitivity in sociometric networks." Working Paper 73-19. Center for Demography and Ecology, University of Wisconsin, Madison.
- Heider, Fritz
1958 *The Psychology of Interpersonal Relations*. New York: Wiley.
- Holland, Paul and S. Leinhardt
1970 "A method for detecting structure in sociometric data." *American Journal of Sociology* 75:492-513.
- 1973 "The structural implications of measurement error in sociometry." *Journal of Mathematical Sociology* 3:85-112.
- Homans, George C.
1950 *The Human Group*. New York: Harcourt, Brace.
- Jencks, Christopher and David Riesman
1968 *The Academic Revolution*. New York: Doubleday.
- Kohlberg, Laurence
1966 "A cognitive-developmental analysis of children's sex-role concepts and attitudes." Pp. 82-173 in Eleanor Maccoby (ed.), *The Development of Sex Differences*. Stanford: Stanford University Press.
- Laosa, L. M. and J. E. Brophy
1972 "Effects of sex and birth order on sex-role development and intelligence among kindergarten children." *Developmental Psychology* 6:409-15.
- Leinhardt, Samuel
1972 "Developmental change in the sentiment structure of children's groups." *American Sociological Review* 37:202-12.
- Lever, Janet
1974 *Games Children Play: Sex Differences and the Development of Role Skills*. Ph.D. dissertation, Department of Sociology, Yale University.
- Maccoby, Eleanor and Carol Jacklin
1974 *The Psychology of Sex Differences*. Stanford: Stanford University Press.
- Mischel, Walter
1970 "Sex-typing and socialization." Pp. 3-72 in Paul Mussen (ed.), *Carmichael's Manual of Child Psychology*. New York: Wiley.
- Omark, D. R. and M. Edelman
1973 "Peer group social interactions from an evolutionary perspective." Paper presented at the Society for Research in Child Development Conference, Philadelphia.
- Parsons, T. W., Jr.
1966 *Ethnic Cleavage in a California School*. Ph.D. dissertation, School of Education, Stanford University. Ann Arbor University Microfilms, No. 56-2602.

- Rivenbark, W. H.
1971 "Self-disclosure patterns among adolescents." *Psychological Reports* 28:35-42.
- Simmel, George
1950 *The Sociology of George Simmel*. Kurt Wolf (ed.). New York: Free Press.
- Sørensen, Aage and Maureen Hallinan
1976 "A stochastic model for change in group structure." *Social Science Research* 5: 43-61.
- Sutton-Smith, B. and B. G. Rosenberg
1971 "Sixty years of historical change in game preferences of American children." Pp. 18-50 in R. E. Herron and Brian Sutton-Smith (eds.), *Child's Play*. New York: Wiley.
- Thibaut, J. W. and H. H. Kelley
1959 *The Social Psychology of Groups*. New York: Wiley.
- Waldrop, M. F. and C. F. Halverson Jr.
1973 "Intensive and extensive peer behavior: longitudinal and cross-sectional analyses." Unpublished manuscript, Child Research Branch, National Institute of Mental Health, Washington, D. C.

APPENDIX A

Conditional Probabilities of Movement from 111D Triad to Other Triad Types for Five Classes by Sex

Class	003	012	102	021D	021U	021C	111D	111U	030T	030C	201	120D	120U	120C	210	300
A Boys (n=10) Girls (n=20)	.0 .025	.0 .161	.0 .397	.0 .012	.0 .054	.036 .070	.0 .054	.0 .054	.143 .017	.0 .004	.036 .058	.0 .021	.179 .083	.036 .008	.429 .029	.143 .008
B Boys (n=12) Girls (n=14)	.015 .041	.088 .190	.118 .252	.015 .007	.029 .102	.074 .061	.132 .027	.132 .027	.059 .054	.015 .0	.103 .068	.044 .034	.059 .116	.029 .014	.132 .020	.088 .014
C Boys (n=10) Girls (n=18)	.0 .016	.167 .091	.139 .176	.0 .010	.0 .085	.083 .049	.083 .029	.083 .029	.028 .055	.0 .0	.083 .153	.028 .020	.139 .231	.111 .020	.083 .052	.056 .013
D Boys (n=23) Girls (n=12)	.011 .032	.069 .190	.109 .159	.040 .016	.025 .048	.065 .032	.105 .032	.105 .032	.055 .032	.084 .0	.145 .190	.044 .0	.109 .159	.062 .016	.138 .063	.018 .032
E Boys (n=13) Girls (n=14)	.038 .048	.094 .075	.302 .370	.047 .0	.019 .055	.019 .015	.038 .034	.038 .034	.028 .0	.0 .0	.151 .144	.009 .007	.189 .137	.009 .021	.057 .062	.0 .034

APPENDIX B

Conditional Probabilities of Movement from 111U Triad to Other Triad Types for Five Classes by Sex

Class	003	012	102	021D	021U	021C	111D	111U	030T	030C	201	120D	120U	120C	210	300
A Boys (n=10) Girls (n=20)	.0 .007	.043 .095	.021 .420	.021 .085	.0 .007	.085 .081	.0 .042	.0 .042	.106 .025	.0 .0	.021 .057	.255 .057	.064 .028	.043 .011	.255 .057	.085 .028
B Boys (n=12) Girls (n=14)	.014 .012	.072 .060	.029 .143	.130 .083	.014 .024	.043 .131	.058 .036	.058 .036	.087 .071	.0 .0	.101 .095	.043 .202	.058 .036	.072 .012	.232 .048	.043 .048
C Boys (n=10) Girls (n=18)	.000 .013	.033 .118	.167 .140	.100 .074	.0 .017	.0 .039	.033 .070	.0 .070	.0 .087	.0 .004	.200 .122	.067 .100	.0 .044	.133 .039	.200 .114	.067 .017
D Boys (n=23) Girls (n=12)	.015 .111	.075 .133	.118 .178	.121 .089	.009 .0	.063 .111	.019 .133	.019 .133	.058 .044	.0 .0	.077 .022	.236 .022	.008 .067	.042 .0	.130 .044	.028 .044
E Boys (n=13) Girls (n=14)	.120 .038	.110 .267	.310 .305	.010 .019	.010 .0	.303 .029	.030 .095	.030 .095	.030 .019	.0 .010	.150 .048	.030 .029	.070 .000	.010 .010	.080 .048	.010 .086

SUBURBAN SOCIAL STATUS: PERSISTENCE OR EVOLUTION?*

AVERY M. GUEST

University of Washington

American Sociological Review 1978, Vol. 43 (April):251-264

This paper focuses on changes in socioeconomic structure of U.S. suburbs between 1920, 1950, and 1970. Persistence in suburban socioeconomic structure has been important in both time periods, but particularly so in the post-World War II period. Moderate evidence of suburban status evolution is found in the period between World War I and 1950. In the earlier time period, evolution was explained by characteristics of both the metropolitan area and the individual community; in the most recent time period, it related primarily to characteristics of the individual community. In both time periods, evolution in suburban socioeconomic structure has been related most strongly to community population growth.

The centrifugal movement of population to the suburbs has been one of the dominant demographic trends of the twentieth century in the United States. Suburban rings were growing faster than their central cities by the early part of the twentieth century, and by the post-World War II period almost all metropolitan population growth was being accommodated in the suburban rings (Kasarda and Redfearn, 1975). While the exact causes of this suburbanization are debated, few would disagree that the mass diffusion of the motor vehicle after World War I, particularly after 1920, has been an important cause and consequence of this trend (Guest, 1973; Tobin, 1976). The high-speed automobile displaced the streetcar as the major means of transportation, permitted the dispersal of the concentrated central population, and the dispersal encouraged the further use of the motor vehicle in the expanded metropolis. While suburban population growth has primarily characterized recent decades, almost half of all 1970 incorporated suburbs with at least 2,500 population were found in the 1900 census (Guest, 1976b). Little is known about how these communities have changed during this massive growth.

Most research on the nature of suburbia has focused on changes in the total population composition of the suburban ring,

especially in reference to the populations of total central cities (for examples, see Schnore, 1965; 1972). As far as we know, only one article in the sociological literature traces changes in individual suburban characteristics over time periods of more than a few years. In that paper, Farley (1964) demonstrated that suburbs in his samples retained an amazingly high persistence in social status between 1920 and 1960 and between 1940 and 1960. In other words, the relative ranking of suburbs in status changed very little over the time periods. Population growth for individual suburbs, Farley's only other independent variable, had only a slight positive effect on status at the second point in time once initial status was considered.

Farley's conclusions were in direct conflict with the general thrust of ecological theory and research which has emphasized the general evolution of community spatial structure. This emphasis is perhaps most evident in the well-known Burgess (1925) hypothesis which argued that, in the twentieth century, central communities became increasingly differentiated from peripheral communities in population and land use characteristics as the metropolis grew in population over long time periods. Older, more centrally located communities experienced such changes as declining social status and shifts away from a predominantly white ethnic composition, while newer, peripheral neighborhoods developed as high status, ethnically white and native-born communities. Such an evolutionary per-

* This research has been supported by HEW Grant 1 R01 HD08961-01 (American Suburban Development). A version of this paper was presented at the April, 1977 annual meeting of the Population Association of America, St. Louis.

spective on areal population composition also is evident in more recent work (Hoover and Vernon, 1962:183-98; Schnore, 1972) which outlines a life cycle pattern for parts of the metropolis.

In his paper, Farley asked how individual suburban communities could primarily persist in status while various research shows that central cities have become increasingly differentiated from their total suburban rings in social status. This differentiation has been primarily due to gains in absolute levels of suburban status, while central city absolute status has remained relatively constant. In explaining this apparent paradox, Farley (1964:44) suggested a model of suburban evolution in which high status suburbs attract the greatest population growth, but high population growth has little effect on further gains in status. High status suburbs show growth in residents with similar high status. This type of model suggests that large gains in absolute suburban ring status thus are due primarily to the increasing share of population found in the high status suburbs, not intrasuburban areal evolution. Subareal social status remains relatively constant in the metropolis, but the population shifts location between low and high status areas. While Farley investigated the empirical effects of population growth on status at a second point in time, he did not actually show the relationship of population growth to initial status.

While Farley's paper was an important landmark in the ecology literature, the questions raised by his work deserve more attention. His study of suburban persistence between 1920 and 1960 involved a sample of only 131 suburbs, found in the 24 largest Urbanized Areas (UA) in 1960 and containing a population of 10,000 in 1920. Furthermore, his 1940 to 1960 comparisons focused on suburbs in only three UAs, Boston, Chicago and Cleveland.

This paper pursues some of the issues raised by Farley's analysis by investigating changes in social status between 1920, 1950, and 1970 for a larger, more representative sample of suburbs. Our basic sample will be drawn from the 3,284 census-defined suburban communities with at least 2,500 residents which were found in

the 243 Standard Metropolitan Statistical Areas (SMSAs) in 1970. Suburbs thus consist of all communities within SMSAs which were not identified as central cities by the U.S. Census Bureau.

Evolutionary Theories

We address two interrelated issues. First, how much persistence is found in the socioeconomic status of American suburbs? Is it as high as Farley found in his smaller sample, or has a process of evolution occurred? Second, what are the major determinants of the evolution which has occurred? Two sets of causal variables will be analyzed, characteristics of the individual suburbs and characteristics of the metropolitan area in which the suburb is found. The distinction is similar to that frequently made between individual and contextual effects. We first discuss the importance of the individual characteristics (spatial position within the metropolis, suburban age or period of formation, and population growth). We then turn to a discussion of metropolitan characteristics (metropolitan area age, metropolitan population growth, and central city density).

According to some interpretations of the Burgess hypothesis (Hawley, 1950: 382-404; Schnore, 1972), the outward expansion of the Central Business District (CBD) leads to increased deterioration of housing, congestion, and the presence of undesirable commercial and industrial uses in affected neighborhoods. The distance of the community from the center of the metropolis becomes an important determinant of its propensity for evolution in social status, with relatively central communities most apt to experience land use changes which would drive out higher status residents.

Research on this interpretation, primarily drawn from the post-World War II period, provides contradictory results. For instance, Haggerty (1971) and Hunter (1974) identified what they claimed were clear declines in social status for centrally located neighborhoods in several American central cities during recent decades.¹

¹ Haggerty (1971) studied census tracts in ten cen-

In contrast, Guest (1972) reported little relationship between areal social status and distance from the center in 37 SMSAs in 1950 and 1960, and he found only a very slight tendency for the relationship to change in the hypothesized direction between 1950 and 1960. Guest's results would seem compatible with Farley's view that subarea status within the metropolis is relatively constant, amid changes in the overall status of total central cities to their total suburban rings.

The relevance of the "distance" theory to suburban communities might be limited, for most suburban communities are not close to the CBD. Actually, some evidence (Guest, 1972; Smith, 1970) suggests that the most peripheral suburban communities may be relatively low in status. This may occur because they are beyond the desirable commuting range of many high status workers. Inner suburban communities thus would have a particular tendency to *increase* in status.

Another interpretation of the Burgess theory is that neighborhoods decline in status over time as the housing becomes obsolescent (Alonso, 1964). Presumably, newly constructed housing is more attractive than older housing in terms of building techniques and facilities. As housing ages, it also allegedly undergoes a process of physical deterioration, so it becomes less attractive for habitation by higher status residents. This interpretation would seem to involve an emphasis on the role of community age or period of development in the evolutionary process. In interpreting this theory, Alonso (1964:228) notes, "The suburbs, as time goes by and buildings age, will become available to those of lower income."

Some research suggests only mild support for this theory in recent decades, at least when communities are distinguished

on the basis of their formation periods. After categorizing census tracts into age cohorts by their period of initial building up, Guest (1974) studied changes between 1940 and 1970 in social status for census tracts in the Cleveland SMSA and changes between 1960 and 1970 in 12 SMSAs. He found only very mild evidence of neighborhood evolution for any age cohort in the period 1940 to 1970 for Cleveland.² His analysis of the 12 SMSAs showed little change between 1960 and 1970 in the status of older relative to newer cohorts, although the changes were generally consistent with the theory. Furthermore, it was difficult to discern clear cross-sectional differences in social status across the census tracts when they were differentiated by their period of building up.

Questions might be raised, though, about whether a focus only on community age or period of formation provides a true test of the aging theory. If new housing is not particularly constructed in newly formed communities, the effects of the housing age process on community status may not be discerned. Thus, the theory suggests that the construction of new housing in a community leads high status persons to locate there, while the lack of new housing leads to a relative decline of status. Within this interpretation of the theory, Farley's (1964) findings about the weak (but evident) effect of population growth in altering suburban social status make sense. The prime determinant of the residential population in a community is the number of housing units, and population growth is primarily determined by the creation of new additional housing units. Additional housing units may be created by the subdivision of existing units or the construction of new units. Subdivision of housing units has traditionally occurred on a large-scale basis toward the center of the metropolis, while the construction of new housing units has been much more

tral cities between 1940 and 1960. He argued that the relatively central tracts were declining in status compared to the peripheral tracts, but an inspection of his published data for several of the central cities suggests more cautious conclusions. Hunter (1974) studied changes in social status for Chicago between 1930 and 1960 through factor analysis. Social status was determined by factor scores at each point; since the factor scores were not always based on the same variables or weighting of variables, it is a little difficult to interpret Hunter's results.

² Guest (1974) did find strong evidence that older census tracts had declined in the period before 1940. He suggested that the declines in status were due to their relatively central location. In the pre-1940 period, the expansion of the CBD was more evident, producing a greater tendency for neighborhood evolution due to changing neighborhood land uses.

prevalent on the outskirts. It seems possible, then, that population growth may lead to positive increases in social status because these communities have the greatest amounts of new housing.

Studies of differences in community evolution across metropolitan areas generally have emphasized the role of metropolitan age and population growth and characteristics of the metropolitan center which would drive higher status persons to the metropolitan periphery (see Guest, 1972; Schnore, 1965; 1972). Most studies have focused on crude differences in status across metropolitan areas between central cities and their total suburban rings. As Hawley (1950:381-404) has interpreted the Burgess hypothesis, metropolitan population growth speeded up the central land uses changes which led to the reorganization of metropolitan status patterns on a concentric zone basis. Metropolitan growth intensified the demand for the limited amount of central space, forcing the land use changes, such as congestion and housing deterioration, which generally have been identified with the flight of higher status persons. Growing metropolitan areas, then, would show a particular reorganization of suburban status patterns, while less rapidly growing metropolitan areas would show a slower rate of change.

Most studies of the status distribution of metropolitan dwellers in the post-World War II period show, however, that metropolitan age is a better predictor of cross-sectional central city-suburban status patterns than current population growth or size (Schnore, 1965:212-3). Older metropolitan areas, regardless of their recent population size or growth, seem to have generally higher peripheral status than newer metropolitan areas.

In clarifying this finding, Guest and Nelson (1978) have shown that the relationship between metropolitan age and the peripheral distribution of high status residents primarily appeared in the period between 1920 and 1950. They suggest that older metropolitan areas in this period had developed unattractive centers as a consequence of earlier population growth. Once the high-speed automobile became available to large segments of the popula-

tion after 1920, the higher social strata deserted the centers of older metropolitan areas. This process did not occur to the same extent in newer metropolitan areas. Guest and Nelson also found that increasing suburban ring relative to central status was characteristic of almost all metropolitan areas in the period 1950-1970. They suggested that the relationship between metropolitan age and status distribution, which had developed at an earlier time point, persisted in the post-World War II period, although the general tendency for higher suburban than central city status increased in all metropolitan areas. According to Guest and Nelson, the fact that suburban status increased rather universally suggested that higher status persons were not differentially pushed by varying degrees of congestion and deterioration in different central cities.

Their analysis also showed that metropolitan size and population growth have been related only weakly to increasing suburban ring relative to central city status, regardless of the time period. In the 1920-1950 period, rapidly growing metropolitan areas did show some tendency to have increasing suburban relative to central city status. They interpreted this as basically indicating support for the Burgess hypothesis and also indicating that the Central Business District might still be expanding. In the 1950-1970 period, rapidly growing metropolitan areas actually showed slightly less tendency than other metropolitan areas for suburban ring status to increase relative to the central city. They argued that the development of rapid speed communication such as the automobile had reduced the demand for the metropolitan center, and thus metropolitan areas were no longer responding to population growth in the traditional way. In fact, Guest and Cluett (1975) have shown that most central business districts in the post-World War II period were not expanding in the traditional manner which was suggested by the Burgess hypothesis.

The implication of Guest and Nelson's results is that individual suburbs may have undergone a status evolution between 1920 and 1950 if they were found in older, rapidly growing metropolitan areas with

disagreeable central cities. In the 1950-1970 period, these characteristics would have little effect on the propensity of suburbs to evolve differentially across metropolitan areas.

Our study will investigate persistence and change in suburban status between 1920, 1950, and 1970. We first examine the persistence of incorporated suburbs with data on their social status for all three dates and then suburbs with data on their status in 1950 and 1970. We then turn to an evaluation of the role of individual community and metropolitan area characteristics in affecting changes in community status over these time periods.

The study selected the dates 1950 and 1970, rather than 1940 and 1960 (as done by Farley), for two primary reasons. First, the speed of American suburbanization intensified after 1950, not 1940 (Guest, 1975). The periods between 1920 and 1950 and 1950 and 1970 thus represent varying tempos of suburbanization. The period 1920-1950 also involved a more dramatic change in the transportation orientation of the suburban ring (from streetcar to automobile) than did the post-World War II period (a continuation of the automobile oriented suburbia). Our second reason for selecting the dates is that an analysis of 1970 rather than 1960 data involves a more recent time point. Of course, Farley's study was published before 1970 census data were available.

Persistence, 1920-1970

Farley's study of persistence from 1920 to 1960 was limited in sample size by his measure of 1920 status, the percentage of 16- and 17-year-olds who were attending school (reported only for suburbs of 10,000 or more in 1920). This measure was used in lieu of better measures of social status that appeared only in later censuses (after 1940) such as the percentage of high school graduates among the population, 25 years and over. Farley did not use a similar measure of 1920 status, the percentage of 16- to 20-year-olds attending school, which was available for all suburbs above 2,500 in 1920 and, thus, would permit investigation of a much larger and varied sample of suburbs.

To determine the feasibility of using the percentage of 16- to 20-year-olds attending school as a 1920 status measure, we correlated our 1920 measure with Farley's and with the 1950 and 1970 status measures, the percentage of high school graduates, 25 years and over. Data for 1950 and 1970 were available for suburbs with at least 2,500 population at the time of observation. These correlations were computed over the 172 suburbs (of 10,000 population) in 1920 with all the available data. The two 1920 status measures correlated .94. The percentage of 16- to 20-year-olds attending school correlated about the same with the percentages of high school graduates in 1950 and 1970 (.77 and .72, respectively) as the percentage of 16- to 17-year-olds attending school correlated with the 1950 and 1970 status measures (.79 and .74, respectively). The value of using the percentage of 16- to 20-year-olds in school as a reasonable measure of 1920 status thus would seem to be justified on the basis of its interrelationship with Farley's measure and its similar correlation with subsequent status.

Our basic sample will consist of the 661 incorporated suburbs with data on their population sizes and social status in 1920, 1950, and 1970. These suburbs also had all the necessary data for the multivariate analysis in the next section. Most of the suburbs eliminated by this constraint were found in New England; constant population sizes of SMSAs are difficult to reconstruct, since the basic units of the SMSA definition in this region are based on townships. While one might quibble with the criteria for inclusion in the sample, various other samples (such as those involving unincorporated communities or including New England communities) had little effect on the results, as long as a full range of communities in terms of their population size was included.

For these communities, the 1920 measure of social status correlated .69 with 1950 status and .63 with 1970 status, while the 1950 and 1970 status measures had correlations of .90. We found that 1920 status explained 48% of the variance in 1950 status, while 1950 status explained 80% of the variance in 1970 status. When

we correlated the 1950 and 1970 status measures for the 1,363 incorporated suburbs with the necessary data for those dates, we obtained a similar .87 correlation.

We also should emphasize that the 1920 status can be predicted almost as well by knowing 1970, as 1950, status. This is a consequence, of course, of the high stability of status distributions between 1950 and 1970.

The correlation of 1920 status with later 1950 status is lower than that found by Farley (.81) in his study of a more select group and slightly different measure, but for a longer time period (1920 to 1960). The difference in our result is probably attributable to the broader range in types of suburbs in our sample. Since Farley's suburbs were required to have 10,000 population in 1920, they were characterized by less variation in growth patterns.

To us, the data suggest somewhat less persistence in social status during the 1920-1950 than 1950-1970 period, as indicated by the clear differences in eventual status explained by initial status. This conclusion must be qualified by two facts, though. First, one correlation involves two decades while the other involves three. Correlations would almost naturally become lower over extended periods, but the differences in explained variance seem too great to be attributable simply to differences in time of observation. Second, the 1920 status measure probably has greater measurement error than 1950 status in predicting eventual status. This is because the 1950-1970 correlation is based on exactly the same measure, percentage of high school graduates, while the 1920-1950 comparison is based on different measures. The fact that measurement error is probably not massive is suggested by the very high correlations found between 1920 and later status in some samples. Thus, in his small sample, Farley found a .81 correlation between 1920 and 1960, while, in our small sample of 172 suburbs (discussed above), we found a correlation of .77 between 1920 and 1950.

The importance of persistence versus evolution in the two time periods also may

be evaluated by whether other possible independent variables have different effects in the two time periods. Obviously, if other independent variables are more powerful in the 1920-1950 period, the existence of community evolution in this period would be more strongly documented.

Multivariate Model

We now consider the effects of individual suburban and metropolitan area characteristics in altering suburban social status between 1920 and 1950, and between 1950 and 1970.

Population growth during the period 1920-1950 and 1950-1970 was operationalized in the same manner as Farley (1964:44). He determined the proportion of the population at the second point in time which was accounted for by the population at the first point in time. This measure was then subtracted from one ($1 - \text{Pop}_1/\text{Pop}_2$). Suburban communities with high population growth thus would have observations near one while negative population growth would be indicated by negative numbers. SMSA population growth during each time period was operationalized in a similar manner.

Suburban age was operationalized as the year of incorporation for the community, or the year first reported in the census as an incorporated community. The earliest data of incorporation was set equal to 1850. The U.S. Census (1963: Tables) reports date of incorporation for communities over 10,000 in 1960, but dates of incorporation for other communities had to be determined by tracing them back through various censuses. We found some difficulty in categorizing dates of smaller communities before the census of 1850. The analysis also was run for communities (above 10,000 population in 1960) with earlier available dates of incorporation, and similar results were obtained as we subsequently report.

Spatial position within the metropolis would ideally be determined by distance from the metropolitan Central Business District (CBD). Unfortunately, many metropolitan areas have multiple central

cities, and many others do not have census-designated central business districts. Furthermore, distances often had to be determined from crude census maps. As a result of these problems, we operationalized spatial location as distance from the central city borders. Four categories were determined on the basis of distance to the nearest central city borders: contiguous (coded 0), within 10 miles (1), between 10 and 20 miles (2), more than 20 miles (3).

In order to simplify the analysis, we treat the distance categories as one interval variable. The results were similar to those obtained by defining dummy variables for each distance zone and are much simpler to present. Furthermore, for metropolitan areas where distance from the CBD was ascertainable, we found that our measure of distance from the central city was generally quite highly correlated. Finally, we point out that the use of our measure of distance perhaps may be justified on the grounds that it has been used by respected researchers in the area of suburban communities (Schnore, 1956).

The other two variables in the analysis are the metropolitan area age and the central city population density (a measure of central congestion). Metropolitan age was measured as elsewhere (Schnore, 1965), the census year in which the central city or cities first reached 50,000 population. Density was measured as the ratio of central city population to land area in square miles.

Table 1 shows the zero-order correlations of the six independent variables with the 1920, 1950, and 1970 measures of social status. Aspects of these cross-sectional correlations are worth noting. First, distance from the central city and date of incorporation have been only weakly related to suburban social status at each time point. These findings are thus consistent with other research, particularly by Guest (1972; 1974) on the importance of these variables in differentiating social status of metropolitan subareas. Second, the major correlate of status at each time point is individual community population growth. Third, of the metropolitan area characteristics, population growth is most strongly related to the

cross-sectional distribution of status among suburbs. However, its zero-order effects may be partially due to its strong correlation with the growth of individual suburban communities. Fourth, older metropolitan areas did not differ from newer metropolitan areas in the average 1920 status of their suburbs, but a slight relationship had developed by 1950 which changed very little by 1970.

Five models of status determination for each time period will be tested. Model 1: status at the second point in time is simply a function of status at the first point in time. Model 2: eventual status is a function of both initial status and the three individual suburban characteristics. Model 3: eventual status is a function of initial status and community population growth. Model 4: eventual status is a function of both initial status and metropolitan area characteristics. Model 5: eventual status is a consequence of initial status, individual suburban characteristics, and metropolitan area characteristics. The variance explained by each model, shown in Table 2, allows determination of the relative predictive power of various sets of variables, in addition to initial status (Model 1). We also have shown the standardized and unstandardized regression coefficients of all predictors in Model 5, the most extensive. By comparing explained variances between the models, we may determine the predictive power of various sets of variables, once initial status has explained the maximum amount of variance.

The following major conclusions may be drawn. First, regardless of the model, the various independent variables are much stronger predictors of eventual status in the 1920-1950 than 1950-1970 periods. For instance, in Model 5, the six individual and metropolitan characteristics add only 4.6% variance to that explained alone by 1950 status. In contrast, the six individual and metropolitan characteristics add 14.8% of the variance to that explained alone by 1920 status.³

³ This paper has focused on the percentage of all variance in social status which could be explained by factors other than initial social status. This allows us to determine the extent of measurable evolution. Our

Table 1. Zero-Order Correlations among Measures of Social Status and Community and Metropolitan Area Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Dist=Distance from central city	—	—	—	—	—	—	—	—	—	—	—	—
(2) INCO=Date of incorporation		—	—	—	—	—	—	—	—	—	—	—
(3) Grw _s =Growth variable, 1920-50			—	—	—	—	—	—	—	—	—	—
(4) Grw _{gr} =Growth variable, 1950-70				—	—	—	—	—	—	—	—	—
(5) Age=Year central city reached 50,000					—	—	—	—	—	—	—	—
(6) Smg _s =SMSA growth variable, 1920-50						—	—	—	—	—	—	—
(7) Smg _{gr} =SMSA growth variable, 1950-70							—	—	—	—	—	—
(8) Dens=Central city density per square mile, 1920								—	—	—	—	—
(9) Dens _c =Central city density per square mile, 1950									—	—	—	—
(10) Stat _s =Percent attend school, 16-20, 1920										—	—	—
(11) Stat _{gr} =Percent high school graduates, 1950											—	—
(12) Stat _{gr} =Percent high school graduates, 1970												—
Mean	1.23	1882	.315	.140	1875	.317	.219	10,959	11,421	26.4	38.7	52.9
Standard Deviation	.96	21	.290	.285	36	.167	.148	5,524	5,667	12.0	13.4	12.7

Source: U.S. Bureau of the Census, 1922: Table 1; 1923: Tables 10, 11; 1952a: Tables 34, 38; 1952b: Table 4; 1971: Tables 32, 34; 1973: Tables 40, 42.

Table 2. Models of Status Determination, 1920-1950

1920-1950				1950-1970		
	b	S.E. of b	B	b	S.E. of b	B
(1) Dist	-1.100**	.357	-.079	.511*	.209	.038
(2) Inco	.018	.016	.029	.018	.009	.031
(3) Grw	14.60***	1.54	.316	11.85***	.87	.266
(4) Age	-.038***	.010	-.102	.00047	.00730	.001
(5) Smg	-4.21	2.57	-.052	-11.82***	1.65	-.137
(6) Den	.00026***	.00007	.105	-.00007	.00005	-.032
(7) Stat	.710***	.029	.634	.797***	.016	.840
R ² (Model 5)	.626			.849		
R ² (Model 4)	.550			.804		
R ² (Model 3)	.584			.835		
R ² (Model 2)	.590			.837		
R ² (Model 1)	.478			.803		

* Statistically significant at .05 level, one-tailed F-test.

** Statistically significant at .01 level, one-tailed F-test.

*** Statistically significant at .001 level, one-tailed F-test.

Second, in both time periods, individual suburban characteristics add more variance to initial status than metropolitan area characteristics (Model 2 versus Model 4). In the 1950-1970 period, in fact, metropolitan area characteristics seem to have trivial effects beyond initial status. The effects of metropolitan area characteristics are more impressive in the earlier time period.

Third, in both time periods, the strongest partial regression effects (other than the initial status) are clearly found for individual suburban growth. In fact, Model 3 shows that population growth added the majority of all additional variance in status, beyond initial status. Consistent with Farley's (1964) earlier study, rapidly growing suburbs showed the greatest relative gains in status. We will analyze these effects in more detail.

procedure has involved partitioning the variance in status into the components due to (a) initial status, (b) other measured independent variables, and (c) other factors (including measurement error, unmeasured variables, and simple random variation). It also is possible to determine the multiple partial correlation (Blalock, 1972:458-9), which involves computing the percentage of variance in b and c which is explained by b. This statistic has an ambiguous interpretation because it only indicates the percentage of what might be conceived as the maximum possible evolutionary variance (b and c) which is explained by measured evolutionary variables. The multiple partial correlation in Model 5 was .284 for the 1920-1950 period and .234 for the 1950-1970 period.

Fourth, in the 1920-1950 period, suburban status evolution was clearly selective of metropolitan area types: older metropolitan areas and those with the most dense central cities. The effects of metropolitan growth were not statistically significant at the .01 level (one-tailed F-test), but were slightly negative.

Fifth, in the 1950-1970 period, metropolitan age and density have trivial effects. Metropolitan growth emerges as a significant negative predictor of status evolution. The most rapidly growing metropolitan areas showed the *least* gains in suburban social status. These results for metropolitan area characteristics in the two time periods generally support Guest and Nelson's (1978) analysis of status distributions between central cities and total suburban rings for the same time periods.

Sixth, in both time periods, the effects of individual suburban spatial location and age are very small in absolute terms.⁴ In

⁴ Some literature (Schnore, 1957; 1963) also has investigated the relationship of social status to the nature of the community as a place of residences or workplaces. There is some confusion in the literature whether the focus is on the relationship of status to type of employment or to the relative prevalence of all residences as opposed to all workplaces. In a recent paper, Guest (1976a) has shown that the overall ratio of jobs to residences in a community is not a good cross-sectional predictor of community social status. However, the nature of the type of employment is a good predictor, particularly whether the community is oriented toward employment in man-

the earliest time period, distance has a statistically significant negative effect on status while in the later period, distance has a statistically significant positive effect. That is, in the earliest time period status gains were most characteristic of inner communities, while in the later time period they were most characteristic of peripheral communities.

Why might higher status suburbs have centralized in the pre-World War II but not in the post-World War II period? One explanation is that the commuting range of the metropolis was particularly limited in the post-World War I period. Employment was still concentrated toward the center, and the rather primitive highway system permitted commuting only within a few miles of the central city. As a result, higher status persons would select residential locations near the central city, for they could have the benefits of better housing and environmental conditions compared with the central city and better

commuting advantages compared with the more peripheral suburbs. The particular desirability of the inner suburbs in the 1920–1950 period also is indicated by the fact (shown in Table 1) that population growth was greatest in the inner suburbs. In the post-World War II period, the development of multi-lane, limited access highways further increased the commuting range of the metropolis and thus decreased the relative advantages of higher status persons locating in the inner suburbs. Furthermore, employment decentralized on a large-scale basis, so that spatial proximity to the center of the metropolis became less important.

Population Growth Effects

In contrast to Farley, we have found other variables to be somewhat more important in explaining eventual status, once initial status is controlled, particularly for the earlier period of observation. For instance, Farley (1964:43) found that population growth added only 4.4 percent variance in the 1920–1960 period to that explained by status alone for his sample of all suburbs. The six variables in our analysis added 14.6% variance (more than three times what Farley found) in the shorter time period of 1920–1950. Population growth alone added about 9% variance in the 1920–1950 period, to that explained by initial status.

The importance of population growth in altering community social status in the earlier time period may be seen by directly comparing our results with Farley's. Based on the regression equation in which initial status and population growth predicted eventual status, Farley (1964:44) reported that a suburb with no growth between 1920 and 1960 would be expected to have 46.6% high school graduates while a suburb which tripled in population, but started at the same status level, would have 51.4% high school graduates (a difference of 4.8%). Our results indicate much stronger effects for the 1920–1950 period. Our similar equation suggests that the no-growth suburb in 1950 would have 33.7% high school graduates, while the 300% growth suburb would have 44.2% high school graduates (a difference of

ufacturing (low status) or service-trade activity (high status). Unfortunately, good data on the nature of employment in suburban communities are not available for suburbs in the 1920–1950 period, and only limited data are available for the 1950–1970 period. For 198 incorporated suburban communities (all with at least 10,000 residential population in 1950), we were able to assemble data on the social status of residents in both 1950 and 1970 and two measures of employment type for each time period—the ratios of (a) manufacturing workers employed in the community to total residential population (M/P) and (b) of retail and wholesale trade workers employed in the community to residential population (T/P) (U.S. Bureau of the Census, 1950:Table 2; 1951a:Table 103; 1951b:Table 103; 1952a:Tables 35, 39; 1969a:Table 3; 1969b:Table 4; 1970:Table 4; 1973:Table 106, 117). In 1950, educational status was correlated .03 and $-.42$ with the M/P and T/P ratios; in 1970, status was correlated .15 and $-.29$ with the same ratios for 1970. To determine the effects of these ratios on changes in social status (the focus of this paper), we predicted 1970 social status with 1950 social status, and the M/P and T/P ratios for each time point. In the sample of 198 suburbs, 1950 status alone explained 79.8% of the variance. The addition of the four other variables added .5% variance. None of the partial regression weights for employment character was statistically significant at the .05 level (one-tailed F-test), and the additional variance explained by these variables beyond initial status would seem to be trivial. Since the longitudinal data are available for few communities and the effects of the employment variables on longitudinal status are essentially trivial, we do not emphasize their role in this paper.

10.5%). While a tripling of population may sound unusual, 76 or 11.5% of the 661 communities had their populations at least triple during the 1920–1950 period. During the 1950–1970 period, however, a tripling of population would have much less effect on altering a community's social status. Using a similar equation, we found that a suburb which showed no growth between 1950 and 1970 would be predicted to have 51.7% high school graduates in 1970, while the 300% growth suburb would have 57.4% high school graduates (a difference of 5.7%).

The implication of this finding is that changes in total suburban ring status clearly could have occurred in the 1920–1950 period as a result of internal reorganization within the suburban ring. The results for the 1950–1970 period are more consistent with Farley's view that changes in suburban ring social status are probably not explained primarily by the internal reorganization of patterns, although some might want to argue whether the 1950–1970 effects are trivial.

Two contrasting models of suburban change may be further clarified by showing, in Figure 1, a simple path diagram of the interrelationships between initial status, population growth, and eventual status in each time period. It can be seen that high status communities in both time periods showed particular growth, indicating that the high status communities were gaining an increasing share of the suburban ring population. The path effect is slightly larger for the most recent time period. In the 1950–1970 period, however,

growth had somewhat less efficacy than in the earlier period in altering community status. The results suggest that high status communities, in the earlier time period, were not only gaining an increasing share of the suburban population but were evolving upward in social status as a result of population growth. In the recent time period, high status communities were clearly gaining in their share of the suburban population, but population growth was not strongly altering their status.

Why might population growth affect suburban status? Earlier, we suggested that it might indicate the construction of new housing which would be disproportionately occupied by high status persons. Unfortunately, good data on the long-term construction of new housing in suburbs are difficult to obtain. It is possible, though, to correlate the percentage of all 1970 housing units which were built after 1950 with the community population growth rate, 1950–1970. We found a very high positive correlation (.88) for the 661 communities observed from 1920–1970, suggesting that growth primarily taps the construction of new housing. Furthermore, population growth seems to have little independent effect on community status, beyond the effects of housing age. Thus, the variance in 1970 status explained by 1950 status and population growth was .835; the variance in 1970 status explained by 1950 status and housing age was .831; the variance in 1970 status explained by 1950 status, population growth, and housing age was .836. This interpretation would suggest more

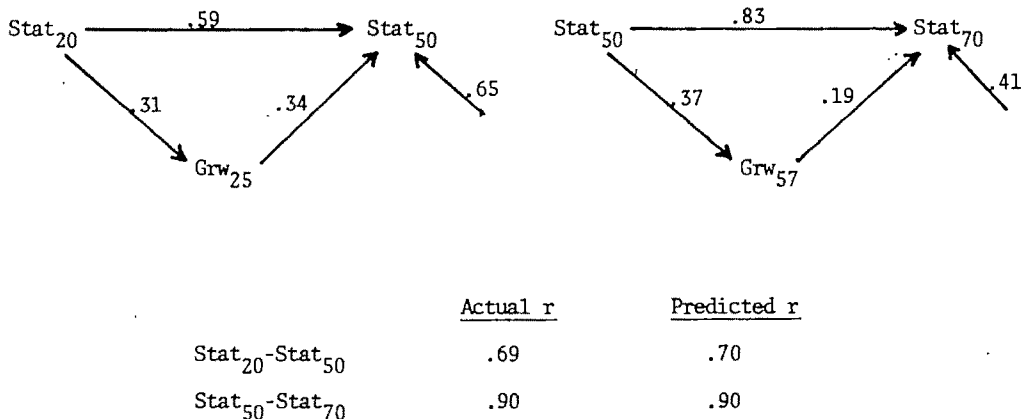


Figure 1. Models of Relationships among Status and Population Growth Variables

support for the aging theory of community evolution than found in the previous section, where we focused on the date of incorporation as an independent variable. It would appear that the construction of new housing attracts higher status persons, but this process occurs somewhat independently of the community's date of incorporation.

Summary and Discussion

Which has been more important—evolution or persistence—in determining the socioeconomic character of contemporary American suburbs? The data suggest two major epochs of suburban status evolution. In the 1920–1950 period, suburbs were differentially changing their status, although persistence was certainly evident. Both metropolitan and individual suburban characteristics were important in this change. Suburbs with gains in status were primarily in older metropolitan areas with dense cities and themselves had high rates of population growth. In the 1950–1970 period, a period of suburban persistence emerged. Amid the massive suburban growth of post-World War II America, most suburbs retained a very high stability of relative ranking in social status. Suburbs with relative gains in social status were growing themselves or their metropolitan areas were failing to grow, but the effects of independent variables were generally small compared to the earlier period.

Overall, the data suggest that community population growth has been the primary variable explaining the evolution of suburbs. As we have pointed out, population growth has been highly correlated with the construction of new housing which has been desirable to high status persons. In fact, high status suburbs have been able to enhance their status relative to other suburbs because they have had the greatest population growth.

Why has the most recent time period been characterized by much more persistence than evolution in social status? An answer to that question is difficult, given the lack of data to test why the temporal change in variable relationships has occurred. It does seem though that the

1920–1950 period represented a more dramatic change in the technological and social organizational character of the metropolis. The transition from the streetcar to the automobile was fairly dramatic. Previous methods of transportation and communication were radically altered, leading to the rise of new valuable suburban centers. High status persons, particularly, moved to these new centers. In contrast, the post-World War II period involved more of the same. The metropolis had already adopted the automobile; the major event was the spread of automobile ownership to even larger numbers of persons, and the even larger flow of persons to the suburban fringe.

The relative persistence of status in the post-1950 period is particularly important in light of the studies which document a growing central city-suburban status gap (Schnore, 1972; Taeuber and Taeuber, 1964). Since period of formation, age and distance variables have generally related weakly to recent status changes and status has been so persistent in the post-World War II period, it seems likely that Farley's model to explain this pattern has a great deal of value. That is, the disproportionate gains in suburban ring status are heavily due to population growth flowing disproportionately to the high status suburbs, thus enhancing their share of the suburban ring population. Thus, suburban rings may evolve upward in social status without significant reorganization of areal patterns within the ring. This does not mean that individual suburbs have not changed their statuses relative to each other, but internal differentiation within the suburban ring would seem to be a highly incomplete explanation of overall status level changes in suburban rings.

This paper, by necessity, has concentrated on the suburbs with available data to study status changes between 1920 and 1970. Obviously, the development of new suburbs since 1920 may be an important means by which the suburban ring can raise its overall level of social status. Since the initial status of the new suburbs is unknown, it is impossible to determine their relative evolution. Dramatic gains or changes in the status of the new suburbs seem unlikely, however. When all 1970

suburbs were cross-classified by date of incorporation and 1970 educational level, we found only a slight tendency for the suburbs incorporated after 1920 (most not in the previous analysis) to be higher in status than the other suburbs. The major contrast was between the suburbs incorporated after 1900 and suburbs incorporated before 1900. As noted previously, date of community formation or incorporation is simply not a good predictor of social status level.

REFERENCES

- Alonso, William
1964 "The historic and the structural theories of urban form: their implications for urban renewal." *Land Economics* 40:227-31.
- Blalock, Hubert M., Jr.
1972 *Social Statistics*. 2nd ed. New York: McGraw-Hill.
- Burgess, Ernest
1925 "The growth of the city." Pp. 47-62 in Robert E. Park, Ernest W. Burgess and Roderick D. McKenzie (eds.), *The City*. Chicago: University of Chicago Press.
- Farley, Reynolds
1964 "Suburban persistence." *American Sociological Review* 29:38-47.
- Guest, Avery M.
1972 "Urban history, population densities, and higher status residential location." *Economic Geography* 48:375-87.
1973 "Urban growth and population densities." *Demography* 10:53-69.
1974 "Neighborhood life cycles and social status." *Economic Geography* 50:228-43.
1975 "Population suburbanization in American metropolitan areas, 1940-1970." *Geographical Analysis* 7:267-83.
1976a "Another look at suburbs and satellites." Unpublished paper, Center for Studies in Demography and Ecology, University of Washington, Seattle.
1976b "The formation and growth of American suburbs." Unpublished paper, Center for Studies in Demography and Ecology, University of Washington, Seattle.
- Guest, Avery M. and Christopher Cluett
1974 "Metropolitan retail nucleation." *Demography* 11:453-507.
- Guest, Avery M. and George H. Nelson
1978 "Central city-suburban status differences: fifty years of change." *Sociological Quarterly* 19. In press.
- Haggerty, Lee J.
1971 "Another look at the Burgess hypothesis: time as an important variable." *American Journal of Sociology* 76:1084-93.
- Hawley, Amos H.
1950 *Human Ecology*. New York: Ronald Press.
- Hoover, Edgar M., and Raymond Vernon
1962 *Anatomy of a Metropolis*. Garden City, N.Y.: Anchor Books.
- Hunter, Albert
1974 "Community change: a stochastic analysis of Chicago's local communities, 1930-60." *American Journal of Sociology* 79:923-47.
- Kasarda, John O. and George V. Redfearn
1975 "Differential patterns of city and suburban growth in the United States." *Journal of Urban History* 2:43-66.
- Schnore, Leo F.
1956 "The functions of metropolitan suburbs." *American Journal of Sociology* 61:453-8.
1957 "Satellites and suburbs." *Social Forces* 36:121-7.
1963 "The social and economic characteristics of American suburbs." *Sociological Quarterly* 4:123-34.
1965 *The Urban Scene*. New York: Free Press.
1972 *Class and Race in Cities and Suburbs*. Chicago: Markham.
- Smith, Joel
1970 "Another look at socioeconomic status distributions in urbanized areas." *Urban Affairs Quarterly* 5:422-53.
- Taeuber, Karl E. and Alma F. Taeuber
1964 "White migration and socio-economic differences between cities and suburbs." *American Sociological Review* 29:718-29.
- Tobin, Gary
1976 "Suburbanization and the development of motor transportation." Pp. 95-111 in Berry Schwartz (ed.), *The Changing Face of the Suburbs*. Chicago: University of Chicago Press.
- U.S. Bureau of the Census
1922 *Financial Statistics of Cities That Have a Population of over 30,000: 1921*. Washington, D.C.: U.S. Government Printing Office.
1923 *Population: 1920, Vol. 3*. Washington, D.C.: U.S. Government Printing Office.
1950 *U.S. Census of Manufacturers: 1947, Vol. 3. Report for Each State*. Washington, D.C.: U.S. Government Printing Office.
1951a *U.S. Census of Business: 1948, Vol. 3. Retail Trade; Area Statistics for Each State*. Washington, D.C.: U.S. Government Printing Office.
1951b *U.S. Census of Business: 1948, Vol. 5. Wholesale Trade; Area Statistics for Each State*. Washington, D.C.: U.S. Government Printing Office.
1952a *U.S. Census of Population: 1950, Vol. 2. Characteristics of the Population, Part for Each State*. Washington, D.C.: U.S. Government Printing Office.
1952b *County and City Data Book: 1952*. Washington, D.C.: U.S. Government Printing Office.
1963 *U.S. Census of Population: 1960, Vol. 1. Characteristics of the Population, Part for Each State*. Washington, D.C.: U.S. Government Printing Office.

- 1969a Census of Business: 1967. Retail Trade. BC67; Report for Each State. Washington, D.C.: U.S. Government Printing Office.
- 1969b Census of Business: 1967. Wholesale Trade. BC67; Report for Each State. Washington, D.C.: U.S. Government Printing Office.
- 1970 Census of Manufacturers: 1967. Area Series:MC67 (3); Report for Each State. Washington, D.C.: U.S. Government Printing Office.
- 1971 U.S. Census of Population: 1970. Number of Inhabitants. Final Report PC(1)-A1. United States Summary. Washington, D.C.: U.S. Government Printing Office.
- 1973 Census of Population: 1970, Vol. 1. Characteristics of the Population. Part for Each State. Washington, D.C.: U.S. Government Printing Office.

A THIRD INTERPRETATION FOR THE GENERATING PROCESS OF THE NEGATIVE BINOMIAL DISTRIBUTION*

WILLIAM W. EATON
McGill University

ANDRÉE FORTIN
University of Montreal

American Sociological Review 1978, Vol. 43 (April):264-267

The two traditional interpretations for the generating process of the negative binomial distribution are reviewed and a third interpretation added. The new interpretation is that spells of time exist during which one or more episodes can occur. This interpretation relies on work by Quenouille formally linking the Poisson, logarithmic, and negative binomial distributions. A method for choosing between the three interpretations with over-time correlations is presented and applied to data on hospitalization for schizophrenia. The longitudinal analysis indicates that the spells interpretation is the correct one.

The distribution of numbers of episodes of given behaviors in a population has often closely conformed to the negative binomial distribution. Traditional derivations of the negative binomial have been shown to model two important social processes. It is traditionally thought of as generated by either a process of reinforcement contagion or a Poisson process operating in a heterogeneous population (Coleman, 1964). We wish to explain a third important formalization of the generating process, present a method for choosing among the three with longitudinal data, and provide an example from data on mental hospitalization.

The frequency of individuals with k episodes of some behavior can be described by the negative binomial in its simplest form:

$$= \binom{\alpha + k - 1}{k} p^\alpha q^k, \quad k = 1, 2, 3, \dots$$

Each of the two traditional interpretations has parameters which are involved in the derivation of the interpretation and which are simple functions of α and q . In the heterogeneity interpretation, the relevant parameters are those of the gamma distribution which indicate the amount of heterogeneity present (Spilerman, 1970). In the contagious process, the relevant parameters are the coefficients of innovation and reinforcement (Coleman, 1964: 301):

It is not generally recognized that there is a third interpretation for the generating process of the negative binomial distribution. In this interpretation, individuals pass through spells of time in which one or more episodes are likely to occur. Between spells, no episodes occur. The spells interpretation was first used by Cresswell and Frogatt (1963) with the Neyman A distribution to describe the pattern of bus driver accidents. It can also be applied to the negative binomial. Quenouille (1949:163), in commenting on studies of bacterial counts, has shown that when "colony counts followed Poisson's

* This research was supported by NIMH Grant MH25691 to William W. Eaton. Thanks are due to Mohinder Massand for typing the manuscript.

distribution [and] the numbers of bacteria per colony were logarithmically distributed[,] . . . the bacterial counts were distributed in the negative binomial form." He includes a proof of the relationship between the three distributions. Although Quenouille's statement is a spatial one, it can also be conceptualized through time in terms of spells. The Poisson and logarithmic distributions have been used many times in describing biological phenomena (for the Poisson, see Feller, 1968:159-64; for the logarithmic, see Williams, 1947). For the Poisson, the number of individuals with n spells is given by:

$$P(n) = \frac{\lambda^n e^{-\lambda}}{n!}.$$

For the logarithmic, the probability of having s episodes in any one spell is:

$$P(s) = \frac{\beta X^s}{s}.$$

An example is episodes of hospitalization for schizophrenia, which may be thought of as occurring during spells. An empirical distribution of episodes has been fit to the negative binomial yielding parameters $q = .56$ and $\alpha = .82$ (Eaton, 1974). Values for the logarithmic and Poisson are calculated from the parameters of the negative binomial, as in Quenouille (1949):

$$X = q = .56;$$

$$\beta = 1/\ln(1-X) = 1.20;$$

$$\lambda = \alpha/\beta = .68.$$

In this example of schizophrenia, the number of spells would be Poisson dis-

Table 1. Predicted Distribution of Spells of Schizophrenia

Numbers of Spells	Predicted Number of Persons
0	—
1	1617
2	551
3	122
4	21
5	3
6	—
	2314

tributed with parameter λ (Table 1), and the number of episodes of schizophrenia during each spell would be logarithmically distributed, with parameters X and β (Table 2). In the case of the spells interpretation, the logarithmic distribution of s episodes per spell is different for each number n of spells, defined by the coefficient of t^s in:

$$[-\beta \ln(1-Xt)]^n.$$

It is impossible to choose between the spells, heterogeneity and reinforcement models solely on the basis of the univariate distribution of recurrences. Arbous and Kerrich (1951) present the bivariate negative binomial for distinguishing the two traditional models, but its assumptions are too restrictive and it is not a strong test anyway. Multivariate data can give evidence of heterogeneity, as in Spilerman's (1970) study of racial disturbances, and longitudinal multivariate data can give evidence of reinforcement. At present there is no agreed-upon analysis which provides evidence of the spells interpretation.

Table 2. Distribution of Episodes of Schizophrenia

Numbers of Episodes	Proportion with Given Number of Episodes for Those with:				
	One Spell	Two Spells	Three Spells	Four Spells	Five Spells
1	.68	—	—	—	—
2	.19	.47	—	—	—
3	.07	.26	.31	—	—
4	.03	.13	.26	.21	—
5	.01	.07	.17	.24	.14
6	.01	.04	.10	.19	.19
7	.00	.01	.06	.13	.19

The best way to choose among the three interpretations is to make use of over-time correlational data. If one divides the period under study into two time periods, all three models predict that the number of episodes during one period should be positively correlated with the number of episodes in the second period. However, for three or more time periods, the three models make distinct predictions.

Under the heterogeneity model, individuals possess an unchanging characteristic which influences their tendency to have episodes. Those who have episodes during one time period are more likely to have episodes in a later time period. More importantly, there is no reason to suspect the correlation to change as the interval between periods increases.

With the positive reinforcement model, individuals having episodes during one time period are reinforced and should be expected to have more episodes later, yielding a positive correlation. As time passes, however, population heterogeneity increases, and thus the between-interval correlations should become stronger later in the period under study. This argument depends to some extent on the nature of the reinforcement process: as Taibleson (1974) has observed, it does not hold for a time-damped reinforcement process.

With the spells interpretation, individuals in a given spell are more likely to have episodes than those not in a spell, again yielding a positive correlation. However, since spells are of limited duration, the correlation should decline over time. The expected pattern is thus opposite to the reinforcement model: the longer the interval between time periods, the lower the correlation.

The pattern of correlations for the data on hospitalization for schizophrenia is consistent with the spells interpretation. For instance, for the cohort of 2,316 schizophrenics admitted in 1961-1962, the number of episodes in 1963 correlated .23 with the number in 1964; .19 with the number in 1965; .11 with the number in 1966; and .09 with the number in 1967 (Eaton, 1974: Table 3). The original author was incorrect in tentatively choosing the heterogeneity interpretation.

If the spells interpretation is correct, what does that imply for the study of schizophrenia? Most individuals in this population have just one spell in this seven-year period (Table 1): 1,617 out of 2,314, or 70%. There is a strong tendency to have just one episode of hospitalization in each spell (Table 2): of those with one spell, 68% had just one episode; of those with two spells, 47% had two episodes; and so forth. Clinicians characteristically monitor the schizophrenic's behavior closely before making treatment decisions such as discharge or discontinuation of drug therapy. Presumably intensive therapy should continue throughout a given spell, and the recurrence of episodes during a spell indicates the decision to discontinue therapy was premature, or perhaps that the individual failed to comply with the treatment regimen. In these cases the chemotherapy may hide symptoms well enough so that the clinician feels the spell is over. If a diagnostic test that is not disrupted by chemotherapy becomes available, the therapist could make these decisions more accurately. As well, the idea that the spells are Poisson distributed could be tested. This interpretation of episodes of schizophrenia is the best fit to the data at hand, but it remains conjectural and can be tested only after significant advances in research in diverse fields.

This type of longitudinal analysis may be relevant to all sorts of chronic behaviors. Examples are periods of unemployment, clusters of life events, criminal actions, political activities, behaviors involving biological rhythms, and many others. The form of analysis usually taken for the study of recurrent behaviors is to study the relationship of individual characteristics to the number of recurrences. Often analysts are proud to account for 5%, 10%, or 15% of the variance in their dependent variables, ascribing the large residual to random influences. The purpose of this paper is to show how random and longitudinal processes can produce variation in the number of recurrences without a necessary relation to any individual characteristics. The longitudinal processes may be the subject of important social theories.

The availability of formalizations does not restrict social life to conform to them. One is not forced into choosing one of these three interpretations for a given set of data. For most human processes of this type the best description may be an inelegant mixture of the processes discussed here, or perhaps a process without a convenient formula. Still, these formalizations alert one to basic processes that may be at work. And, when they fit the data well, they can be extremely useful.

REFERENCES

- Arbous, A.G., and J.E. Kerrich
1951 "Accident statistics and the concept of accident proneness." *Biometrics* 7:340-423.
- Coleman, S.
1964 *Introduction to Mathematical Sociology*. New York: Free Press.
- Cresswell, W. L., and P. Frogatt
1963 *The Causation of Bus Driver Accidents*. London: Oxford University Press.
- Eaton, W.
1974 "Mental hospitalization as a reinforcement process." *American Sociological Review* 39:252-60.
- Feller, W.
1968 *An Introduction to Probability Theory and Its Applications*. 3rd ed. New York: Wiley.
- Quenouille, M.H.
1949 "A relation between the logarithmic, Poisson, and negative binomial series." *Biometrics* 5:162-4.
- Spilerman, S.
1970 "The causes of racial disturbances: a comparison of alternative explanations." *American Sociological Review* 35:627-49.
- Taibleson, M. H.
1974 "Distinguishing between contagion, heterogeneity, and randomness in stochastic models." *American Sociological Review* 39:877-80.
- Williams, C. B.
1947 "The logarithmic series and its application to biological problems." *Journal of Ecology* 34:253-72.

ERRATA

In the article "Multiple Roles and Role Strain" (ASR December, 1977), "on which extraneous valuables" (p. 930, col. 2, 12 lines from the bottom) should read:

on which no extraneous valuables have been staked. . . .

THOMAS J. FARARO's name was inadvertently omitted in the 1977 list of editorial consultants (ASR December, 1977).

In the comment on Inverarity (ASR August, 1977), equation (3) was incorrect. The correct equation is:

$$\begin{aligned} \underline{y} &= \underline{\beta}(\underline{\alpha}\underline{x} + \epsilon) + \underline{u} \\ &= \underline{\Pi}' \underline{x} + \underline{v}. \end{aligned}$$

COMMENTS

THE MEANING OF THE IQ-DELINQUENCY RELATIONSHIP

(COMMENT ON HIRSCHI AND
HINDELANG, ASR AUGUST, 1977)

Hirschi and Hindelang (1977) review several studies which demonstrate a strong relationship between IQ and delinquency, and suggest that the ideological commitments of the sociology of deviance have prevented sociologists working in this area from recognizing this association. It is doubtful that sociologists are as ignorant of this association as Hirschi and Hindelang would have us believe. Indeed, in the course of their paper they cite the work of many sociologists who have offered various interpretations of this relationship. Be this as it may, the provocative aspect of their article consists of the way they attempt to account for this association. They argue that the data show IQ scores to be measures of stable abilities which are largely independent of cultural or social influences.¹ They then come up with a neogenetic theory of delinquency, and suggest that IQ affects the likelihood of delinquent behavior through its effect upon school performance. Their model holds that low IQ leads to difficulties in school and thereby enhances the probability of delinquent behavior. It is our contention that Hirschi and Hindelang are blatantly wrong in their interpretation of IQ scores. For whatever reason, they ignore a large body of literature concerning the impact of environmental factors upon IQ scores.

Hirschi and Hindelang report two studies (West, 1973; Wolfgang et al., 1972) which show that controlling for attitudes toward or trouble in school reduces the relationship between IQ and delinquency to insignificance. Viewing IQ scores as indicators of stable, innate ability, they conclude that these studies suggest the mechanism operating in the relationship between IQ and delinquency: "IQ affects the likelihood of delinquent behavior through its effect on school performance" (p. 584). But, if one assumes that IQ scores are unstable and subject to social influences, it might reasonably

be argued that the causal ordering is of a different sort. The delinquent's negative attitudes toward school or his teacher's negative attitude toward him may lead to troubles at school and to a lack of motivation to develop the abilities tapped by IQ tests. There are several models of this type which might be formulated. What they all have in common is that they view IQ as somehow following from the causes of delinquency or from delinquency itself, rather than serving to cause delinquent behavior.

Hirschi and Hindelang attempt to rule out such models by marshalling evidence to the effect that IQ scores are not subject to social influences. However, their evidence consists mainly of the strong test-retest correlation which has been shown to exist between IQ measurements separated by several years. They discount Rosenthal and Jacobson's (1968) study which shows a relationship between teacher expectations and student IQs by noting that the study received negative reviews by Snow (1969) and Thorndike (1968). Unfortunately, they fail to present the content of the criticisms made of the study, or to mention that Rosenthal has attempted to address the criticisms made by these two reviewers (Rosenthal, 1969; Rosenthal and Rubin, 1970). However, regardless of the status of the Rosenthal and Jacobs (1968) study, there is a profusion of evidence indicating that IQ scores are responsive to social environmental factors.

First, consider the following data summarized by Bronfenbrenner (1975). He analyzed data reported in several studies of identical twins raised apart and found that for pairs with the same number of years of schooling the average IQ difference was 1.45. For pairs who differed in schooling by five years or more the difference was 10.4. The correlation of Binet IQ scores for twins raised in the same town was .83; for those brought up in different towns the correlation was .67. For twins attending the same school in the same town the correlation was .87, while for those attending schools in different towns the coefficient was .66. When raised by relatives the correlation was .82, but when raised by unrelated persons the figure was .63. Finally, separated twins living in a similar community displayed a correlation of .86; the coefficient for those growing up in dissimilar localities was .26. Thus, the less similar the environment in which identical

¹ Note that accepting Hirschi and Hindelang's contentions concerning IQ tests would demand that sociologists view blacks as less intelligent than whites, and lower-class individuals as less intelligent than upper-class persons.

twins are raised, the less the correspondence in their IQs. These data clearly show that modifications in the social environment can produce changes in IQ scores.

A second, even more convincing, body of data comes from programs designed to increase the IQ scores of low social class and minority group individuals. Bereiter (1969; Bereiter and Engelmann, 1966) in four replications of his remedial approach has produced IQ gains averaging 15 points. Klaus and Gray (1968) reported increases averaging nine points on the Stanford-Binet, and the increase was still evident three years after training had terminated. Karnes et al. (1968; 1970) produced Stanford-Binet IQ gains of 12 to 14 points over the academic year. Blank and Solomon (1968) increased Stanford-Binet scores by an average of 14.5 points. Heber et al. (1972) found that their Milwaukee project resulted in a mean IQ of 124 for the experimental group, and 94 for the control group—a difference of 30 points. The sizes of these increases are interesting given Hirschi and Hindelang's observation that studies show delinquents to have an IQ which averages about nine points less than that of the conformers. In the studies just cited, the researchers were able to produce changes that were as great, or in most cases substantially greater, than this nine point difference. And, most of the programs produced these changes in about a year's time.

Given these studies, it is evident that Hirschi and Hindelang are presenting an outdated view of IQ, a view which is inconsistent with the empirical evidence. Experts in the area of intelligence no longer view IQ as a global mental ability which one inherits. Rather, it is viewed as a broad set of verbal and problem-solving skills which are better labeled academic aptitude or scholastic readiness. The contemporary view of what IQ tests measure is described by Whimbey and Whimbey (1975:67,6-7), researchers in the area of intelligence:

Intelligence is taken to be, in large part at least, a habitual approach to problem solving—a learned mental skill. Training this skill is accomplished through demonstration and guided practice. . . . There is a common misconception that the Stanford-Binet measures innate mental capacity, whereas academic aptitude tests such as the SAT are more heavily dependent on learning. This is obviously impossible, since the items in the adult Stanford-Binet and the items in the SAT (especially the Verbal Section) are almost identical. . . . Thus, for example, in her book *Psychological Testing*, Anne Anastasi discusses the SAT in the chapter entitled "Group Tests of General Intelligence." The Graduate Record Exam (GRE) and the Law School Admission Test (LSAT) are basically upward extensions of the SAT. They measure

primarily the same abilities, but at a higher level of difficulty. . . . The differences between the tests are small in comparison with their similarities.

As has been widely publicized, college freshmen have demonstrated lower SAT scores in recent years. Given the equivalence of the SAT and the Stanford-Binet, Hirschi and Hindelang would be forced to draw the conclusion that college students are becoming less intelligent. This is an unlikely proposition, given that the demographic characteristics of college freshmen have not changed much since the early 1970s.

In keeping with this learned-skill perspective on IQ, many school administrators are substituting reading comprehension tests for IQ tests. The score distributions obtained are almost identical with IQ tests. This is not surprising as questions appearing in reading comprehension tests are almost indistinguishable from those appearing in standard verbal intelligence tests (Farr, 1970). Given that many educators now view IQ tests and reading tests as measuring roughly the same thing, one more study might be cited. This study is important because it not only addresses the issue of the instability of IQ, but also provides some clues in regard to the causal sequence which may be operating.

James Jones (1965), Kenneth Clark and the staff of the Haryou Planning Project studied two kinds of schools—middle-class black schools and lower-class black schools. They found that in the early grades (second, third, and fourth) there was no significant difference in reading comprehension scores between the children in the two kinds of schools. In the fifth and sixth grades, the children from the lower-class schools began to fall behind. By the eighth grade there were large differences between the two groups. These findings, which were replicated by Ryan (1965), suggest that lower-class kids do not come to school with low IQs which cause them to fail (and perhaps turn to delinquency), but rather something about their interaction with the school system seems to stagnate their growth. The data are consistent with a model where lower-class children encounter trouble with the school, which leads to failure to learn and lower IQ scores.

Finally, even if IQ tests did measure some innate, stable mental capacity, there is the problem of the respondent's motivation while taking the test. IQ tests conducted by the school system are almost always administered in large groups. The delinquent is often described as an unmotivated student who does little school work and receives failing grades. Hirschi and Hindelang accept this description

of the delinquent and make it a part of their model; i.e., low intelligence leads to low performance and negative attitudes toward school which, in turn, enhances the likelihood of delinquency. But, if these students are not motivated to do academic work on any other day of the school year, why should they be motivated to perform to the best of their ability on the day the IQ tests are administered? Clearly, there is no reason. As a result the IQ scores of persons who frequently participate in delinquent acts should be viewed as suspect, and any relationship between IQ and delinquency becomes meaningless.

In summary, Hirschi and Hindelang are probably correct in contending that there is a relationship between IQ scores and delinquency. However, the issue is the meaning that is to be attributed to this relationship, and Hirschi and Hindelang contribute little to the topic other than obfuscation. Indeed, should their naive view of IQ be taken seriously, the field of sociology would be taking a giant step backwards.

Ronald L. Simons
Iowa State University

REFERENCES

- Bereiter, C.
1969 "The future of individual differences." *Harvard Educational Review* 39:310-3.
- Bereiter, C., and S. Engelmann
1966 *Teaching Disadvantaged Children in Preschool*. Englewood Cliffs: Prentice-Hall.
- Blank, M., and F. Solomon
1968 "A tutorial language program to develop abstract thinking in socially disadvantaged preschool children." *Child Development* 39:379-89.
- Bronfenbrenner, Urie
1975 "Nature with nurture: a reinterpretation of the evidence." Pp. 114-44 in Ashley Montague (ed.), *Race and IQ*. New York: Oxford University Press.
- Farr, R.
1970 "Measuring reading comprehension: an historical perspective." Pp. 187-97 in F. Green (ed.), *Twentieth Yearbook of the National Reading Conference*. Milwaukee: National Reading Conference.
- Heber, R., H. Garber, S. Harrington, C. Hoffman, and C. Falender
1972 *Rehabilitation of Families at Risk for Mental Retardation: Progress Report*. Rehabilitation Research and Training Center in Mental Retardation, University of Wisconsin, Madison.
- Hirschi, Travis and Michael J. Hindelang
1977 "Intelligence and delinquency: a revisionist review." *American Sociological Review* 42:571-87.
- Jones, James A.
1965 "Education in depressed areas: a research sociologist's views." Paper at the Fourth Work Conference on Curriculum and Teaching in Depressed Areas, Teachers College, Columbia University, New York.
- Karnes, M. B., A. S. Hodgins, R. L. Stoneburner, and W. M. Studley
1968 "Effects of a highly structured program of language development on intellectual functioning and psycholinguistic development of culturally disadvantaged three-year-olds." *Journal of Special Education* 2: 405-12.
- Karnes, M. B., J. A. Teska, and A. S. Hodgins
1970 "The successful implementation of a highly specific preschool instructional program by paraprofessional teachers." *Journal of Special Education* 4:69-80.
- Klaus, R. A., and S. W. Gray
1968 "The early training project for disadvantaged children: a special report after five years." *Monographs for the Society for Research in Child Development* 33, No. 4.
- Rosenthal, R.
1969 "Empirical vs. decreed validation of clocks and tests." *American Education Research Journal* 6:689-91.
- Rosenthal, R. and Lenore Jacobson
1968 *Pygmalion in the Classroom*. New York: Holt.
- Rosenthal, R. and D. Rubin
1970 "Pygmalion reaffirmed." Pp. 139-55 in J. Elashoff and R. Snow (eds.), *Pygmalion Reconsidered*. Worthington, Oh.: Jones.
- Ryan, William
1965 "Racial imbalance, cultural deprivation, and reading achievement." Mimeo. Advisory Committee on Racial Imbalance and Education, Background Paper No. 7, Massachusetts State Board of Education.
- Snow, R.
1969 "Unfinished Pygmalion." *Contemporary Psychology* 14:197-9.
- Thorndike, R. L.
1968 "Review of R. Rosenthal and L. Jacobson, 'Pygmalion in the Classroom'." *American Education Research Journal* 5:703-11.
- West, D. J.
1973 *Who Becomes Delinquent?* London: Heinemann.
- Whimbey, Arthur and Linda Shaw Whimbey
1975 *Intelligence Can Be Taught*. New York: Dutton.
- Wolfgang, Marvin, Robert M. Figlio and Thorstein Sellin
1972 *Delinquency as a Birth Cohort*. Chicago: University of Chicago Press.

INTERNAL POLITICAL ORGANIZATION AND THE WORLD ECONOMY OF INCOME INEQUALITY

(COMMENT ON RUBINSON,
ASR AUGUST, 1976)

Recent research on the comparative study of income inequality has suggested that gains in economic development often have little bearing on the stratification system of a nation (Vandendries, 1974; Fishlow, 1972; United Nations, 1967; 1971; Stack, 1976). In light of this work the developmental model of income de-stratification put forth by such writers as Kerr and his colleagues (1964) and Kuznets (1963) needs to be, at the very least, supplemented by other models. Rubinson's (1976) recent work presents a fruitful alternative to the developmental model. The strategic position of a nation in world economy in terms of state strength, direct control over production in other states, and relative independence from the external market, is linked to several internal processes such as occupation differentiation, economic expansion, and class formation that directly shape the internal stratification system. Rubinson's command of the evidence and his application of Wallerstein's (1974) propositions on world economy to stratification analysis are masterful. World-economy indicators such as exports/GDP and debits on investment income are found to have a significant impact on income inequality within nations independent of the level of economic development. Level of development is found to have no statistically significant relationship with the level of overall income inequality. Rubinson effectively demonstrates the superiority of his model over the developmental model.

However, Rubinson does not test his model against the other major alternative, the political model. Writers such as Lenski (1966), Cutright (1967), and Peters (1973) stress the internal distribution of power in their theories of income stratification. The presence and intensity of democratic political institutions, the degree of equality in the distribution of political power, the degree of working-class political participation, and other factors are viewed as effective means for reducing income inequality through state action such as minimum wage laws, full employment policy, and progressive income taxation. While Rubinson criticizes Lenski for deleting world-economy factors from his model, Rubinson might be criticized for omitting indicators of the internal distribution of political power.

We anticipate that the inclusion of a measure of internal political organization would im-

prove Rubinson's work. We have reanalyzed Rubinson's data and have included a measure of internal politics, Jackman's (1975:64-5) index of democratic performance. This index consists of four elements which measure political competition, political representativeness, free access to information, and electoral participation. Nations with the highest scores would have higher voter turnouts, a multiparty system, strong minority parties, regular elections, and a free press. Under these conditions the elite is most likely to be under effective pressure for reducing income inequality (Lenski, 1966; Cutright, 1967; Jackman, 1974; 1975).

Table 1 presents the results of the regression analysis involving our indicators of level of development, political organization, and world economy on the overall degree of income inequality. Our measure of income inequality, the Gini index, and our data source, Paukert (1973), are the same as Rubinson's. Data on the index of democratic performance were available for 36 of the 47 nations in Rubinson's study. The level of development is measured in terms of GDP/capita, which was the indicator of development (the others being energy consumption/capita, the log of GDP/capita, and the log of energy consumption/capita) and which provided the best fit for the curve. Our index of a nation's dependence and involvement in world economy is exports/GDP, which is the best theoretical and empirically confirmed indicator available from Rubinson's model.¹

The most important determinant of the degree of inequality is the index of democratic performance ($\beta = .530$). Controlling for level of development and exports/GDP, the greater the democratic performance the less the income inequality. The index of world eco-

¹ For our indicator of the dependence and the extent of involvement in world economy we used exports/GDP as opposed to the seven indicators analyzed by Rubinson. Only three of the indicators used by Rubinson were significantly related to the degree of overall inequality at the .05 level. These were exports/GDP, imports/GDP, and government revenue/GDP. Exports/GDP was a better predictor of overall inequality than imports/GDP, so the former was used. We did not utilize government revenue/GDP for theoretical reasons. This index is not an accurate measure of state strength in world economy. State strength needs to be measured by a different, perhaps qualitative variable in order to measure the nature of the key qualities of state strength in world economy, such as tariff and trade policies, and the rules and regulations controlling production within a nation's boundaries (see Rubinson, 1976:642, for a discussion of the qualities underlying state strength).

Table 1. The Effects of Democratic Performance, Exports/GDP, and Level of Economic Development on the Degree of Inequality in the Distribution of Income (N=36)

Variable	Regression Coefficient	Standard Error of Coefficient	Computed Value of t	Beta Coefficient
Index of Democratic Performance	-.003	.001	-2.338*	-.445
Exports/GDP	.215	.108	1.993*	.323
GDP/capita	-.000	.000	-.735	-.132
R ² = .29				

* Statistically significant at the .05 level.

nomy is also a significant predictor of inequality at the .05 level, even with the other variables controlled. As predicted, the greater the involvement in world economy the greater the inequality. As in Robinson's study, we find that the level of economic development is not related significantly to inequality once we control for the other variables. The sign is in the expected relation, but the size of the t-statistic indicates an insignificant relationship.

As in the case of prior research, given that our data are cross-sectional, the direction of causality is open to question. For example, it is customary to assume that equality sustains democracy just as we assume that democracy nurtures equality. For this reason our results should be viewed with some caution.

While exports/GDP in the present study and four out of eight indicators of world-economy variables in Robinson's paper have proven their importance in shaping the degree of income inequality, the internal political model should not be neglected. Elementary indicators of a nation's internal political organization, as the present paper suggests, are important predictors of inequality of incomes.

Steven Stack
Alma College

REFERENCES

- Cutright, Phillips
1967 "Inequality: a cross-national analysis." *American Sociological Review* 32:562-78.
- Fishlow, A.
1972 "Brazilian size distribution of income." *American Economic Review* 62:391-402.
- Jackman, Robert
1974 "Political democracy and social equality." *American Sociological Review* 39:29-45.
1975 *Politics and Social Equality*. New York: Wiley.
- Kerr, Clark, J. T. Dunlop, Frederick Harbison and Charles Myers
1964 *Industrialism and Industrial Man*. New York: Oxford University Press.
- Kuznets, Simon
1963 "Quantitative aspects of the economic growth of nations, VIII: the distribution of income by size." *Economic Development and Cultural Change* 11:1-80.
- Lenski, Gerhard
1966 *Power and Privilege*. New York: McGraw-Hill.
- Paukert, Felix
1973 "Income distribution at different levels of development." *International Labor Review* 108:97-125.
- Peters, B. Guy
1973 "Income inequality in Sweden and the United Kingdom: a longitudinal analysis." *Acta Sociologica* 16:108-20.
- Robinson, Richard
1976 "The world-economy and the distribution of income within states: a cross national study." *American Sociological Review* 41:638-59.
- Stack, Steven
1976 *Inequality in Industrial Society: Income Distribution in Capitalist and Socialist Nations*. Ph.D. dissertation, Department of Sociology, University of Connecticut.
- United Nations
1967 *Economic Survey of Europe in 1965, Part II: Incomes in Postwar Europe: A Study of Policies, Growth, and Distribution*. Geneva.
1971 *Income Distribution in Latin America*. New York.
- Vandendries, Rene
1974 "Income distribution in Peru after World War II." *Journal of Developing Areas* 8:421-36.
- Wallerstein, Immanuel
1974 *The Modern World System: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century*. New York: Academic Press.

DETERRENCE AND SOCIAL CONTROL*

(COMMENT ON SILBERMAN, ASR JUNE, 1976)

The appearance of Silberman's (1976) "Toward a Theory of Criminal Deterrence" in the ASR represents a significant step in the

*We would like to acknowledge the helpful suggestions from Theodore Anderson, Pat Lauderdale, and Richard Hall on an earlier draft.

study of social control. Four years earlier the ASR published William Goode's (1972) presidential address to the ASA in which he criticized sociologists for their failure to develop a model of force and force-threat as a mechanism of social control. In a tone reminiscent of Dennis Wrong's (1961:188) classic comment that sociologists have ignored "the degree to which conformity is frequently the result of coercion rather than conviction," Goode (1972:508) complained that the "systematic study of force as a distinct phenomenon or set of processes has been singularly neglected" in the discipline. Ironically, a number of sociologists (as well as economists, political scientists, and criminologists), under the rubric of "deterrence theory," have been developing an empirically grounded theory of the role of force-threat (i.e., threats of physical and material deprivation emanating from the State) in the social control process. Quantitative evidence began to appear in the mid-1960s, and by 1974, Thomas and Williams (1974) were able to compile a bibliography of 225 items concerning the deterrent effect of sanctions.¹

Unfortunately, this rapidly growing interest in deterrence theory has not yet been reflected in the discipline's leading journals. In his effort to encourage sociologists to study force, Goode (1972) failed to mention the deterrence research. More recently, Janowitz (1975), after reviewing the state of the field of social control theory, concluded by urging that the role of force be considered more seriously in sociological research without acknowledging the existence of the deterrence literature.

Silberman's (1976) article is the first piece of deterrence research to appear in either the ASR or AJS.² Because the deterrence perspective does address a crucial issue in sociology and because this article will be the first exposure to that perspective for many sociologists, the adequacy of Silberman's presentation and tests of the deterrence propositions must be examined carefully.

Although most deterrence research has employed official crime statistics with states or counties as the units of analysis, the need for survey research with individuals as the units of analysis has been recognized repeatedly. Silberman's is one of only a few studies using the survey method to examine relationships be-

tween self-reported law violation and the perceived certainty and severity of punishment. While the previous studies using individuals as the units of analysis involved self-reports of one or two offenses, or self-reports of general delinquency, Silberman obtained self-reports of violations for eleven offenses and the perceived certainty and severity of legal punishment for each of these among a sample of college students. For each of the three variables (Criminal Involvement, Certainty, and Severity), he summed over the eleven offenses and used these composite measures to test his hypotheses.

Silberman's general strategy, which is consistent with the current approach in deterrence research, is to examine the conditions under which the threat of punishment is a deterrent—i.e., the conditions under which a perceived high threat of legal punishment is associated with low criminal involvement. In his introduction he extracts from the literature several hypotheses concerning conditional relationships and claims that his research is a test of these hypotheses. Unfortunately, his effort contains several errors, inconsistencies, and omissions so that parts of the product are a misrepresentation of the current state of deterrence theory and research. Of central concern is the effect of these errors on the interpretations given to the roles of perceived severity and perceived certainty of legal punishment in the social control process. After examining Silberman's treatment of these issues, we will turn to several of the more serious methodological errors which further confound his analysis.

MUTUAL EFFECTS OF CERTAINTY AND SEVERITY OF PUNISHMENT ON CRIMINAL INVOLVEMENT

Silberman's statements concerning the mutual effects of Certainty and Severity on Criminal Involvement ignore all the issues raised in previous research. In his introduction he observes that previous research indicates a "negative association between certainty of punishment and crime rates, and no association between severity of punishment and crime rates except for homicide" (p. 443). Crime statistics for states do reveal this pattern of zero-order correlations, but Silberman fails to note that for nearly all major felonies there is a substantial negative zero-order correlation between certainty and severity among aggregates. Logan (1972) discovered that when certainty (conviction rate) is controlled, severity (median sentence) has a weak to moderate negative partial correlation with the crime rate for nearly all major felonies. At the aggregate

¹ For an early statement of some of the theoretical issues, see Chambliss (1967). For a recent summary of the research see Andenaes (1975).

² Since the publication of Silberman's research, two other articles concerning deterrence have appeared in the ASR (Meier and Johnson, 1977; Erickson et al., 1977).

level, certainty appears to be a suppressor variable in the relationship between severity and crime rate. This finding, that aggregates which have severe penalties tend to have low conviction rates, has been of interest to researchers (see Bailey and Smith, 1972). Ross (1976) has proposed a hypothesis which he calls the "neutralization of severe sanctions" to explain this kind of phenomenon. He argues that where the penalty becomes so severe that it exceeds what is considered fair by the population, reactions occur in the criminal justice process which lead to a reduction in the conviction rate. For example, prosecutors will be more likely to agree to plea bargaining efforts, and juries will be less likely to find the defendant guilty. The salient policy implication is that an increase in the severity of punishment could reduce the certainty of punishment so much that the crime rate might increase.

There is no reason to assume, however, that this aggregate pattern of relationships is replicated at the individual level. In fact, Silberman reports a slight positive correlation between his measures of perceived Certainty and Severity ($r = +.15$). A different argument, which might be called the "credibility of severe sanctions" hypothesis, has been offered at the individual level in previous writings (see Tittle and Logan, 1973:384), but Silberman does not test it. This hypothesis states that perceived Severity of punishment has a deterrent effect only when the perceived Certainty of punishment is high. In other words, Certainty is a conditional variable in the relationship between Severity and Criminal Involvement. For example, a person who believes that the penalty for using marijuana would be very severe if he were arrested, but who believes there is virtually no possibility of being arrested, is not likely to be deterred by the threat of a severe penalty which he thinks will never be administered. But among those who do believe there is some possibility of being arrested, those who believe the penalty would be severe are less likely to commit the offense than those who believe the penalty would be mild. Since the publication of Silberman's article, Teevan (1976) has presented evidence supporting this hypothesis.

Silberman fails to test the credibility hypothesis at the individual level. Instead, he presents a significant negative zero-order correlation between Certainty and Criminal Involvement and a nonsignificant positive zero-order correlation between Severity and Criminal Involvement. He concludes on this basis plus partial regression coefficients—without even examining an interaction model—that Severity does not have a deterrent effect.

In addition to his failure to test this crucial hypothesis, Silberman's measure of perceived Severity of legal punishment seems invalid, and any conclusions he reaches concerning the role of this variable in the social control process must be considered with extreme caution. His measure is whether or not the respondent believed he would receive the state's maximum penalty if convicted of committing the offense. But there is evidence from other research (California Assembly, 1968) that people vary in what they believe is the legal maximum penalty for particular offenses, and Silberman's measure does not take into account this variation. In his analysis, a person who believes the state's maximum penalty for using marijuana is a \$100 fine and believes he would receive the maximum penalty scores just as high on the Severity index as a person who believes the maximum is a six month jail sentence and believes he would receive the maximum. But even if everyone did know the legal maximum penalty, Silberman's measure would be inadequate. A person who believes he would not receive the maximum but, nevertheless, would receive a relatively severe penalty would not score higher on Silberman's scale than a person who believes he would receive no penalty at all.

MUTUAL EFFECTS OF CERTAINTY, PEER INVOLVEMENT, AND MORALITY ON CRIMINAL INVOLVEMENT

Peer Involvement as a Conditional Variable in the Relationship between Certainty and Criminal Involvement

In his introduction, Silberman suggests that the extent of involvement with peers who engage in criminal behavior (Peer Involvement) might affect the relationship between Certainty and Criminal Involvement. He offers the hypothesis that a high level of involvement with such peers (measured as the number of the eleven offenses for which the respondent knows someone who was arrested) "may reduce the effect of threat of punishment" (p. 443), but offers no rationale for the hypothesis. In fact, three hypotheses concerning characteristics of peers as a conditional variable have appeared in the deterrence literature, and each of these focuses on the role of threat of social disapproval from peers as a mechanism of social control.

One hypothesis (see Jensen, 1969:196-7) suggests that the threat of legal punishment is a significant deterrent only when peers would react by informally punishing (via social disap-

proval or stigmatization) a person who is legally punished for violating a law. The informal sanction contingent upon formal, legal punishment is seen as the real deterrent, and the only function of legal punishment in the social control process is the threat it presents of exposing an offender to his peers. But in the absence of a threat of social disapproval contingent upon legal punishment, the threat of legal punishment should not have a deterrent effect on criminal behavior. Among people whose peers are involved in criminal behavior and thus would not express social disapproval if the person were exposed as an offender, there should be no relationship between perceived Certainty of legal punishment and Criminal Involvement. This hypothesis is derived from differential association theory and appears to be the one Silberman is advocating.

The second hypothesis (see Jensen, 1969: 196-7) is derived from Matza's (1964) discussion of drifters and predicts a pattern of statistical interaction opposite from the one Silberman predicts. Matza claims that as long as a person has conventional ties (i.e., ties to people who support the law), the threat of social disapproval from these people is the source of conformity to the law. A person with strong conventional ties will conform even if he perceives the Certainty of legal punishment to be low because, although his law violation probably would not be detected by the police, it might be detected by his peers who would disapprove of the violation. Only when conventional ties have been severed does the fear of legal punishment per se become the major obstacle to criminal involvement. Thus, the hypothesis predicts that the perceived Certainty of legal punishment will not be related to Criminal Involvement among those respondents whose friends are *not* criminally involved. Such people conform to the law regardless of the perceived Certainty of legal punishment.

The third hypothesis is an additive model and represents Wrong's (1961) argument that both the avoidance of physical and material punishment and the avoidance of social disapproval are sources of motivation to conform to norms. According to this hypothesis, Peer Involvement (translated into Threat of Social Disapproval) and perceived Certainty of legal punishment should have additive effects on Criminal Involvement. The magnitude of the effect of Certainty should not depend on the level of Peer Involvement.

Silberman's test of the interaction of Certainty and Peer Involvement is confounded by a simultaneous control for Morality (discussed later), but the appropriate figures can be reconstructed from the numbers in his Table 4a (p.

450) and are presented in Table 1. Unfortunately, the data required for significance tests are not presented in the article. These figures indicate that high (above the mean) perceived Certainty is accompanied by low Criminal Involvement (number of offenses the respondent has committed) regardless of the level of Peer Involvement. Furthermore, low Peer Involvement (i.e., high threat of social disapproval) is accompanied by low Criminal Involvement regardless of the level of Certainty. Silberman's results support the additive hypothesis and, thus, are consistent with the three other empirical studies which have addressed this issue in deterrence research (Jensen, 1969; Burkett and Jensen, 1975; Grasmick and Appleton, 1977).

Table 1. Mean Score on Criminal Involvement within Categories of Peer Involvement and Certainty

	High Peer Involvement	Low Peer Involvement
High Certainty	3.43	2.72
Low Certainty	4.21	3.23

Morality as a Conditional Variable in the Relationship between Certainty and Criminal Involvement

Silberman's major contribution to deterrence research is his finding concerning the interaction of moral commitment to the law and threat of legal punishment in the deterrence process. Drawing on the theoretical writings of Zimring and Hawkins (1973) and others, he predicts that the threat of legal punishment has a deterrent effect on criminal behavior only under the condition of low moral commitment to the law. High moral commitment indicates that people have internalized the law and will conform even if they perceive a low threat of legal punishment. The only previous test of this hypothesis considered level of moral commitment to be a property of a law, rather than of an individual; but Silberman correctly argues that moral commitment is a characteristic of people. There is no law which everyone has internalized and no law which no one has internalized. Therefore, he measures moral commitment to the law (Morality) as a variable characteristic of people in the sample.

Again, however, Silberman's test of the proposed conditional relationship is confounded by the simultaneous control for another variable. But the appropriate numbers can be extracted from his Table 4a and are presented in Table 2. Among those respondents who score high on the Morality index, perceived

Certainty of legal punishment has no deterrent effect: there is no difference in the number of offenses reported between those who perceive a high Certainty and those who perceive a low Certainty. But among those who have not internalized the law (i.e., people in the Low Morality category), people who perceive the Certainty of legal punishment to be high report, on the average, having committed 3.64 of the offenses, while those who perceive the Certainty to be low report an average of 4.65 of the offenses. On the other hand, Morality has an effect on Criminal Involvement regardless of the level of perceived Certainty. In both the high and the low Certainty categories, those who score low on the Morality index have committed noticeably more of the offenses than those who score high on the Morality index.

Table 2. Mean Score on Criminal Involvement within Categories of Morality and Certainty

	High Morality	Low Morality
High Certainty	2.45	3.64
Low Certainty	2.37	4.65

The importance of this finding should not be overlooked. It reveals that the basic proposition of deterrence theory applies only to part of the population—those with low levels of internalization of the law. Any research which does not take into account this conditional relationship will underestimate the magnitude of the deterrent effect of threat of legal punishment within that segment of the population for which the threat of punishment is a deterrent.

ERRORS IN THE USE OF PATH ANALYSIS

In addition to the problems discussed above, Silberman commits several serious errors in the use of path analysis. In fact, given the pattern of relationships he uncovers and his interaction hypotheses, he should not have chosen to present the combined effects of the independent variables on Criminal Involvement in the form of a path analysis. He claims that his correlation matrix "suggests that the four independent variables [Severity, Certainty, Morality, Peer Involvement] fit a causal chain model" (p. 448). Despite the fact that correlation matrices never suggest causal chains, and despite the fact that his path diagram (which actually is a single-equation model in which all four independent variables are treated as exogenous) does not depict the causal chain he describes (p. 449), Silberman at

least should have noted that path analysis assumes that all effects of independent variables are additive. His own tables demonstrate that the data do not meet this assumption (i.e., the interaction effect of Certainty and Morality, as well as the interaction effect of Peer Involvement and Morality discussed below). In fact, his entire argument in the introduction consists of hypotheses about interaction effects.

In the latter section of his analysis, Silberman switches to nine of the eleven offenses as the units of analysis (rather than individuals) and performs another path analysis. The variables are the Percent Committing the Offense and the sample means for Severity, Certainty, Morality, and Peer Involvement for each offense. Silberman's errors at this level are blatant. First, in the offense-level path diagram (p. 452), Morality suddenly becomes an intervening variable between Certainty and Percent Committing the Offense, while in the individual-level diagram both Morality and Certainty are presented as exogenous. But Silberman offers no rationale for this change. Second, although he appears to recognize that partial standardized regression coefficients based on four independent variables for only nine cases might not be very meaningful, he does not hesitate to draw substantive conclusions from these figures in his discussion. Third, he continues to discuss interaction even though his method of analysis assumes additivity. Finally, he is oblivious to a glaring case of multicollinearity. At the offense level, the zero-order correlation between Morality and Certainty is $+0.91$, and both of these variables have strong negative correlations of approximately equal magnitude with Percent Committing the Offense (-0.90 and -0.82 , respectively). Despite the near perfect correlation between Morality and Certainty at this level, he enters these two variables in the same equation for Percent Committing the Offense, and concludes that only Morality has a direct effect and that all the effect of Certainty is through Morality as an intervening variable. But the difference between the correlation of -0.90 between Morality and Percent Committing the Offense, and the correlation of -0.82 between Certainty and Percent Committing the Offense is not significant. With only nine cases, the difference most likely is an artifact of sampling error. Had the relative magnitudes been reversed so that the correlation was slightly higher for Certainty than for Morality, the path analysis would have shown that Certainty had a strong direct effect on Percent Committing the Offense and that the correlation between Morality and Percent Committing the Offense was spurious due to the effects of

Certainty on these two variables (i.e., no direct effect for Morality). Much of Silberman's discussion toward the end of his article is based on his evidence that, at the offense level, Morality is an intervening variable between Certainty and Percent Committing the Offense. But the evidence is not strong enough to justify his conclusions.

A MODEL OF SOCIAL CONTROL:
LEGAL PUNISHMENT, SOCIAL DISAPPROVAL,
AND
INTERNALIZATION OF THE LAW

Over fifteen years ago, Dennis Wrong (1961) urged sociologists to consider three sources of motivation to conform in their theories and research about social control: (1) the internalization of norms, (2) the threat of social disapproval from peers, and (3) the threat of physical and material sanctions. While the early deterrence literature focused on one aspect of the threat of physical and material sanctions (certainty and severity of legal punishment), more recent research, including Silberman's, has attempted to incorporate the other two sources of motivation. Buried under the problems noted above in Silberman's article is a model of social control which incorporates threat of legal punishment, threat of social disapproval, and the internalization of norms.

In the model, internalization of norms has a primacy effect. As described above, Silberman's data suggest that people who are morally committed to the law are more likely to conform than people who are not morally committed, and that among the morally committed group the perception of relatively certain legal punishment is not an additional deterrent. A further manipulation of figures in Silberman's Table 4a (presented in Table 3) suggests that among those who are morally committed, involvement with peers who engage in criminal behavior has hardly any effect on the tendency to violate laws. But the figures also reveal that among those who have *not* internalized the law, a high threat of social disapproval (i.e., low Peer Involvement) does appear to be a significant mechanism of social control. Previously, Table 2 indicated that the perception of a high certainty of legal punishment is a source of conformity, but (as with the threat of social disapproval) only among people having a low level of moral commitment to the law. Finally, tables presented here and Silberman's own Tables 4a and 4b reveal that the effects of certainty of legal punishment and threat of social disapproval on conformity are additive and that these effects occur only when moral commitment is low.

Table 3. Mean Score on Criminal Involvement within Categories of Morality and Peer Involvement

	High Morality	Low Morality
Low Peer Involvement	2.24	3.61
High Peer Involvement	2.62	4.97

Silberman's data, therefore, suggest that one segment of the population conform because they have internalized the law. Their conformity is independent of threats of legal punishment or social disapproval. But among those who have not internalized the law, the threat of social disapproval and the perceived certainty of legal punishment are independent sources of motivation to conform. Furthermore, had Silberman measured perceived severity of legal punishment properly and performed the appropriate analysis, he probably would have found that among those who believe the certainty of legal punishment is relatively high, the perception of a severe penalty if apprehended is an additional source of conformity to the law.

Silberman is the first researcher to have considered all these variables simultaneously, and, thus, he has made a significant contribution to deterrence research. Moreover, he has introduced many sociologists to the literature concerning this set of variables, and we hope his article will have a long-range impact on sociologists' efforts to study the threat of force as a mechanism of social control. Our comments have been critical, not because we think the article is unimportant, but because we think it is extremely important. Our only concern has been to clarify the pattern of relationships among these variables which has been emerging in the deterrence literature to which Silberman has contributed.

Harold G. Grasmick
University of Oklahoma
Steven D. McLaughlin
Utah State University

REFERENCES

- Andenaes, J.
1975 "General prevention revisited: research and policy implications." *Journal of Criminal Law and Criminology* 66:338-65.
- Bailey, W. and R. Smith
1972 "Punishment: its severity and certainty." *Journal of Criminal Law, Criminology and Police Science* 63:530-9.
- Burkett, S. and E. Jensen
1975 "Conventional ties, peer influence, and the

fear of apprehension: a study of adolescent marijuana use." *Sociological Quarterly* 16:522-33.

California Assembly

1968 *Deterrence Effects of Criminal Sanctions*. Sacramento: Assembly of California.

Chambliss, W.

1967 "Types of deviance and the effectiveness of legal sanctions." *Wisconsin Law Review* 67:703-19.

Erickson, M., J. Gibbs and G. Jensen

1977 "The deterrence doctrine and the perceived certainty of legal punishments." *American Sociological Review* 42:305-17.

Goode, W.

1972 "The place of force in human society." *American Sociological Review* 37:507-19.

Grasmick, H. and L. Appleton

1977 "Legal punishment and social stigma: a comparison of two deterrence models." *Social Science Quarterly* 58:15-28.

Janowitz, M.

1975 "Sociological theory and social control." *American Journal of Sociology* 81:82-108.

Jensen, G.

1969 "'Crime doesn't pay': correlates of a shared misunderstanding." *Social Problems* 17:189-201.

Logan, C.

1972 "General deterrent effects of imprisonment." *Social Forces* 51:64-73.

Matza, David

1964 *Delinquency and Drift*. New York: Wiley.

Meier, R. and W. Johnson

1977 "Deterrence as social control: the legal and extralegal production of conformity." *American Sociological Review* 42:292-304.

Ross, H.

1976 "The neutralization of severe penalties." *Law and Society Review* 10:403-13.

Silberman, M.

1976 "Toward a theory of criminal deterrence." *American Sociological Review* 41:442-61.

Teevan, J.

1976 "Deterrent effects of punishment: subjective measures continued." *Canadian Journal of Criminology and Corrections* 18: 152-60.

Thomas, C. and J. Williams

1974 "The deterrent effect of sanctions: a selected bibliography." Mimeo. Metropolitan Criminal Justice Center, College of William and Mary, Williamsburg.

Tittle, C. and C. Logan

1973 "Sanctions and deviance: evidence and remaining questions." *Law and Society Review* 7:371-92.

Wrong, D.

1961 "The oversocialized conception of man in modern sociology." *American Sociological Review* 26:183-93.

Zimring, Frank and Gordon Hawkins

1973 *Deterrence: The Legal Threat in Crime Control*. Chicago: University of Chicago Press.

ATTRIBUTION, GOALS, AND DEVIANCE

(COMMENT ON HARRIS,
ASR FEBRUARY, 1977)

Harris's (1977) recent article on the systematic omission of the sex variable from major theories of deviance is, for the most part, a splendid contribution to deviance theory. We would like to express, however, a pair of demurrers, relative to his treatment of strain theories and his introduction of the concept of self-attributions.

Harris argues that structured strain theory is counterfactual when faced with the sex variable because women have very low crime rates despite the fact that they massively have been denied access to legitimate economic opportunities. As Harris interprets strain theory, such a denial of access should give rise to higher crime rates of women. However, such counterfactuality only arises if one assumes that women have the same goals as men. Clearly, they do not. Occupational achievement is vastly higher in the hierarchy of goals of men than of women, while the latter have traditionally placed a greater emphasis on domestic goals. Crudely, men are interested in money while women, until very recently, have been more interested in love and babies. While readily admitting variation in goal emphasis by social class, structured strain theory would have it that deviance arises due to a discrepancy between means and goals, *whatever those particular goals may be*.

Given the differing goals of men and women sketched above, we then ask whether the goals-means discrepancy is greater for men than for women. Our suggested answer is that the discrepancy for males is both greater and lasts longer—perhaps forever—during a lifetime. Since the principal means to the goals of women is marriage, and since approximately 90% of women marry at least once, it would seem that access to domestic goals is, at least initially, available to almost all women. Also, the attainment of domestic goals occurs much earlier in life than the attainment of economic goals. While marriages, on the average, are established during one's early twenties, the attainment of economic success is delayed typically until one's forties or fifties. Thus, in the late teens and the twenties, the age group for which the crime rates are typically highest, women would see most of their domestic goals achieved. However, for men of this age group, although their domestic goals may be achieved or achievable, economic goal attainment appears either impossible or is located in a distant future. For such young men, jobs providing

immediate and substantial economic gratification are not forthcoming. Hence, there occurs a high economic frustration level and a correspondingly high crime rate among young men but not among women.

The above suggestion of goals differing by sex implies that when women do commit crimes, those crimes should arise disproportionately out of frustrations of domestic goals. Consistent with this implication, Wolfgang (1958:207) and, more recently, Harris (1976) have reported that a much higher percentage of the homicides by women arise within domestic situations than is the case for men. Analysis of the offender-victim relationship shows that women who commit homicides or assaults are more likely to select family members, or lovers, or sex rivals than are men. The above suggestion of sex differences in goals also implies that, due to the linkage of violence and domestic goals, and compared to men, a higher percentage of women's offenses should be violent offenses rather than property offenses. This appears to have been true until about 1962 when women's share of property offenses began to exceed their share of violent offenses (Simon, 1975:38-9). The relative increase in women's share of property offenses is, of course, attributable to the increased importance of economic goals among women.

That the attainment of men's economic goals is often delayed until late in life implies that, as men either approach those goals or redefine at about age forty their goals and circumstances into ones of success (Tausky and Dubin, 1965), their crime rates should decline. All extant data show such a decline in male crime rates with age (Uniform Crime Reports, 1971:120). However, the implications of our above suggestions for the relationship of age and crime among women are less clear. Although marriage is the principal means for the attainment of women's domestic goals, it also provides opportunities for their frustration. To some extent the woman may find the gains of marriage, home, and children illusory. Or, through marriage a woman may "inherit" her husband's goal-frustration. Also, since the advent of economic activity (such as labor force participation) is later in life for women than for men, opportunities for economic frustration and crime also may come later. English data suggest that women's criminal behavior comes somewhat later in life than does men's (Simon, 1975:96).

We may further argue that, until recently, women's economic goals, in addition to their domestic goals, have been more readily attainable than those of men. For women whose goals have been principally domestic, the attainment of those goals through marriage has

also constituted the attainment of their economic goals. Through marriage the woman traditionally has acquired a "property right" in the husband's income. Such a merging of the legitimate means to both the economic and domestic goals of women into the form of a husband has meant that the means to satisfy women's economic goals have been as widely attainable as the means to domestic goals.

We suggest that, through recognition of the different goals of men and women, structured strain theory can readily adapt to the introduction of the sex variable. Such a recognition suggests that, as women's economic goals have escalated in recent decades, their patterns of economic deviance should begin to approximate those of men. As Adler (1975:15) notes, "By every indicator available, female criminals appear to be surpassing males in the rate of increase for almost every major crime." Further, Simon's (1975:36-9) data show that the marked increase in crimes by women is almost entirely due to an increase in property offenses. Meanwhile, the rates of violent crime among women, including homicide, have remained stable. This is precisely the pattern of deviance our interpretation of structured strain theory would predict.

Finally, we turn to Harris's categorizing of certain types of theories—labelling and containment theories—as "self-attribution" theories. He argues that such theories are not subject to the same criticisms as goal-oriented theories. Essentially, self-attribution theory proposes that goals are considered in terms of whether or not they are attributable to the individual's perception of her/himself (the self-concept). Thus males and females are "type-scripted" into the pursuit of different goals. This seems to simply sneak goals in the theoretical backdoor after they have been cast out the front.

Joseph Harry
Northern Illinois University
Mary C. Sengstock
Wayne State University

REFERENCES

- Adler, Freda
1975 *Sisters in Crime*. New York: McGraw-Hill.
- Harris, Anthony
1977 "Sex and theories of deviance." *American Sociological Review* 42:3-16.
- Harris, Richard
1976 "Urban homicide: Detroit." Paper presented at the meeting of the Western Association of Sociology and Anthropology, Calgary.

- Simon, Rita
1975 Women and Crime. Lexington: Lexington Books.
- Tausky, Curt and Robert Dubin
1965 "Career anchorage: managerial mobility motivations." American Sociological Review 30:725-35.
- U.S. Department of Justice
1971 Uniform Crime Reports. Washington, D.C.: U.S. Government Printing Office.
- Wolfgang, Marvin
1958 Patterns in Criminal Homicide. New York: Wiley.

ITEMS (Continued)

■ **LAWRENCE E. COHEN** (Determinants of Juvenile Court Dispositions) is Assistant Professor of Sociology at the University of Illinois, Urbana. He is engaged (with Kenneth Land and Marcus Felson) in research which attempts to forecast crime rate trends and to build dynamic macrosocial indicator models of the criminal justice system. **JAMES R. KLUEGEL** is Assistant Professor of Sociology at the University of California, Riverside. He is pursuing research on recruitment to authority positions in organizations as an aspect of socioeconomic achievement.

■ **JEFFREY C. ALEXANDER** (Formal and Substantive Voluntarism) is Assistant Professor of Sociology at UCLA. He is working on a study of Watergate as a crisis in civil society and of the symbolic and structural requisites for an effective and substantively voluntary response to social crisis. He is author of *Theoretical Logic in Sociology* (2 Vols., University of California Press, in press).

■ **JAMES R. LINCOLN** (Community Structure and Industrial Conflict) is Assistant Professor of Sociology at the University of Southern California. He is involved in studies of Japanese business organizations in southern California and interorganizational networks of juvenile justice system agencies.

■ **MURRAY WEBSTER, JR.** (Status Generalization) is Professor of Sociology at the University of South Carolina. He is completing a textbook of theoretical social psychology (for Wadsworth), and he has just finished a chapter on roles for the 1978 *Annual Review of Sociology*. His previous books are *Sources of Self-Evaluation* (with Barbara Sobieszek; Wiley, 1974) and *Actions and Actors* (Winthrop, 1975). **JAMES E. DRISKELL, JR.** earned his M.A. degree at the University of South Carolina and is working as a Juvenile Counselor for the State of South Carolina. He continues to study organizing processes and interpersonal behavior from both theoretical and applied perspectives.

■ **DONNA EDER** (Sex Differences in Children's Friendships) is a Ph.D. candidate in Sociology at the University of Wisconsin, Madison. Her research is on the influence of teacher behavior and classroom structure on students' attitudes and achievements. **MAUREEN T. HALLINAN** is Visiting Associate Professor of Sociology at Stanford University. She is doing research on individual and organizational determinants of children's friendships, on the formation structure and stability of children's cliques, and on models of growth in academic achievement. In

1974 she published *The Structure of Positive Sentiment* (Elsevier).

■ **AVERY M. GUEST** (Suburban Social Status) is Associate Professor of Sociology and Associate Director of the Center for Studies in Demography and Ecology at the University of Washington. He is analyzing the ecological development of American suburbs since 1920. His general research focuses on patterns of population distribution within American metropolitan areas.

■ **WILLIAM W. EATON** (Generating the Negative Binomial Distribution) is Associate Professor of Sociology at McGill University. His research interests include the sociology of mental disorders, quantitative models for the study of chronic behaviors, and comparative ethnic relations. **ANDRÉE FORTIN** is a Ph.D. candidate in the Department of Sociology at the University of Montreal. She is doing research on the sociology of development and quantitative methodology.

■ **RONALD L. SIMONS** (Comment on Hirschi and Hindelang, ASR August, 1977) is Assistant Professor of Sociology and Anthropology at Iowa State University. He is involved in research concerned with the role of various types of social stress in the etiology of disease; the impact of the state of the economy upon the incidence of crime; and the various sociological and ethical issues associated with capital punishment.

■ **STEVEN STACK** (Comment on Robinson, ASR August, 1976) is Assistant Professor of Sociology at Alma College. His research includes a development of a synthesis of the developmental, world-economy, and political models of income stratification.

■ **HAROLD G. GRASMICK** (Comment on Silberman, ASR June, 1976) is Associate Professor of Sociology at the University of Oklahoma. He is studying the public's perceptions of the criminal justice system. **STEVEN D. McLAUGHLIN** is Associate Professor of Sociology at Utah State University. He has just completed research on the consequences of the timing of the first birth for the mother's status attainment.

■ **JOSEPH HARRY** (Comment on Harris, ASR February, 1977) is Assistant Professor of Sociology at Northern Illinois University. His research interests include deviance and homosexuality. He is author of *The Social Organization of a Stigmatized Community: The Case of Male Homosexuals* (Praeger, forthcoming). **MARY C. SENGSTOCK** is Associate Professor of Sociology at Wayne State University. Her research centers on ethnicity and the sociology of law.

New Publications Available from ASA...

_____ ***1978 Guide to Graduate Departments of Sociology***

Members & Students: \$4.00

Non-members & Institutions: \$10.00

_____ ***1978 ASA Directory of Members***

Members: \$5.00

Non-members & Institutions: \$10.00

_____ ***1978 Directory of Departments of Sociology***

Members: \$5.00

Non-members & Institutions: \$10.00

_____ ***Journal of Health & Social Behavior Index***

(Volumes 1-17, 1960-1976)

Members: \$5.00

Non-Members & Institutions: \$7.50

To order any of the above, check publication(s) desired and send coupon, enclosing payment, to: American Sociological Association, 1722 N Street, NW, Washington, DC 20036.

(All orders must be prepaid)

Name _____

Address _____

City

State

Zip

AMERICAN SOCIOLOGICAL REVIEW

THE CAUSES AND COST OF RACIAL EXCLUSION FROM JOB AUTHORITY*

JAMES R. KLUEGEL

University of California, Riverside

American Sociological Review 1978, Vol. 43 (June):285-301

Recent studies have shown an improvement in the socioeconomic circumstances of black Americans. Yet a substantial difference remains, raising the question of what accounts for this difference. In this paper we examine the role of job authority (position in the authority hierarchies of organizations) as a source of this persistent discrepancy. Based on the comparison of samples of white (n = 1,088) and black men (n = 625) we find that: (1) the average job authority of black men is markedly lower than that of white men; (2) this discrepancy has its source in a number of factors including organizational characteristics, education, occupational position, and direct within-occupation exclusion from authority; (3) black men receive a lower income return to authority than do whites, with the discrepancy in income return being especially pronounced at higher occupational status levels; and (4) the exclusion from authority is costly; on the average across occupations it accounts for approximately one-third of the total black-white income gap.

Recent studies document an improvement over the last decade in the socioeconomic circumstances of black Americans. The gap between black and white men in education and occupational status has narrowed (Featherman and Hauser, 1976), and the ratio of black to white income has increased (Haworth et al., 1975; Johnson and Sell, 1976). Yet substantial differences in the socioeconomic circumstances of these two groups remain; this raises the question of what accounts for these persistent differences.

A large number of studies within both sociology and economics argue that the

major source of the socioeconomic achievement gap lies in differential experiences within the labor market (Gordon, 1972; Jencks et al., 1972; Masters, 1975). Simply stated, one can divide the set of factors that might account for the black-white difference into two groups: prelabor market and labor market factors. In the category of prelabor market factors we can place educational level, cognitive skills, and other characteristics acquired by individuals before they assume a job. Research estimating the amount of the socioeconomic achievement gap that is attributable to characteristics that black and white men bring to the point of entry into the labor market (Duncan, 1969; Masters, 1975) consistently finds that prelabor market factors account for at most a moderate amount of this difference.

Therefore, to understand what accounts for the racial difference in socioeconomic achievement this research directs one to look at racial differences in both the allocation of jobs and in experiences within a given job. Research by sociologists on the influence of these labor market factors has focused principally on racial differences in the allocation of occupations (Duncan, 1969; Stolzenberg, 1975). Economists

* This research was carried out under grants from the National Science Foundation to Professors David L. Featherman and Robert M. Hauser (G1-31604X and G1-44336) and from the Center for Population Research of the National Institute of Child Health and Human Development to the Center for Demography and Ecology of the University of Wisconsin, Madison (HD-05876). Any opinions, findings, conclusions or recommendations are those of the author and do not necessarily reflect the views of the National Science Foundation. Thanks are offered to David L. Featherman and Robert M. Hauser for the use of the data on which this report is based, and for helpful comments on earlier versions of this work. The valuable comments and suggestions of anonymous reviewers are also thankfully acknowledged.

have examined the impact of a broad range of variables, including characteristics of industries and firms (Rees and Shultz, 1970), unions, housing segregation, migration (Masters, 1975), and the employment stability of jobs (Gordon, 1972).

In spite of the wide range of factors studied in this body of research, one aspect of labor market experience, the holding of positions in the authority hierarchies of organizations (job authority), has been largely ignored in empirical examinations of the sources of black-white differences in socioeconomic achievement. Such research, however, is clearly called for because there are reasons for believing that the racial difference in access to job authority is not only a source of black-white differences, but perhaps represents one of the most important and most persistent sources.

Promotion within authority hierarchies is an aspect of general socioeconomic achievement that seems singularly susceptible to particularistic manipulation. Criteria for promotion, as several observers have noted (Dalton, 1959; Collins, 1971; Gordon, 1972; Kanter, 1975), tend to be vaguely specified; often they involve notions of loyalty, good character and leadership potential. In contrast, the link between educational credentials and the type of occupation held is readily visible and hence open to scrutiny. Racial discrimination based on the exclusion of blacks from occupations for which they hold appropriate credentials as a consequence would seem to be easier to identify and programmatically address than discrimination stemming from the denial of access to authority positions. The very vagueness of criteria for promotion makes them susceptible to subtle manipulation for discriminatory purposes.

There are several ways that the criteria for promotion may work to the disadvantage of blacks. Employers may assume that blacks, because of presumed social and cultural differences, cannot socially interact with whites. This assumption may result in the placement of blacks in only those authority positions that principally involve interaction with other blacks—positions involving either interaction with

black clientele or the supervision of black workers—and thus may restrict the supply of authority positions available (Fernandez, 1975). Both the identification of potential candidates for promotion and the training for authority positions are heavily influenced by informal social interaction in and outside of the workplace. Racial separation in such informal social interaction, done either intentionally or not, can effectively limit the promotion chances of blacks in predominantly white organizations (Ferman, 1968; Fernandez, 1975). Even policies with seemingly benign intent can result in racial exclusion. In this regard several observers (Ferman, 1968; Morgan and Van Dyke, 1970; Fernandez, 1975) point to a prevailing belief among blacks and whites alike that blacks must be overqualified to be promoted to authority positions. Such a policy is often rooted in the desire to be certain that blacks will succeed if promoted.

In addition to the discriminatory use that may be made of criteria for promotion within authority hierarchies, there is a more fundamental aspect of the authority structures of organizations that is of interest for understanding racial differences in socioeconomic achievement. The authority hierarchy of organizations represents a principal source of social control by owners and managers of corporate firms. Concerning social control and discrimination, Thurow (1969:117) speaks of the tendency of the dominant group in a society to interact with the subservient group as a discriminating monopolist or monopsonist; specifically, the discriminator "may want to work with, buy from, or hire Negroes, but he insists on specifying the relationships under which the two parties will meet and how the Negro will respond." One way to maintain this position is to maintain principal control over decision-making positions in organizations; thus, according to Thurow (1969:117), "the discriminator may prefer to hire Negro maids, Negro garbage collectors, or to work with Negroes if he can be in a position of authority." Also, concerning the social control uses of the authority hierarchy, radical economists (Gordon, 1972; Edwards, 1975) have underscored the potential use of promotion

within authority hierarchies as a strategic means of forestalling class-based threats to a capitalistic economic system. Proponents of this viewpoint argue that job authority is a major dimension of labor market segmentation (Edwards, 1975) into primary and secondary labor markets. Jobs within the primary labor market are characterized by career status, providing stability of employment, with increases in wages, status and authority resulting from increased work experience. Jobs within the secondary labor market, in contrast, are characterized by instability of employment, with little opportunity for advancement. Radical economists have argued that the allocation of blacks to the secondary labor market (and derivatively the exclusion from authority) can serve to negate the development of class consciousness among workers who otherwise share a basically similar class position.

If indeed the allocation of job authority is as susceptible to particularistic influence as it appears to be, and if the allocation of authority positions is of such fundamental importance as a means of social control, then we can clearly expect to achieve a better understanding of the sources of racial differences in socioeconomic achievement by researching racial differences in authority attainment. As we noted earlier, however, this topic has been little researched. Although there is a significant body of research that speaks indirectly to this topic—for example, census data and other research (Hauser and Featherman, 1974) document the persistent underrepresentation of black men in managerial occupations and in the self-employed category—there is little systematic analysis directly addressed to the issues of how large a gap in job authority exists, what factors account for this gap, and how important a part the job authority discrepancy plays in accounting for the black-white difference in socioeconomic achievement. Unfortunately, research that has directly addressed these issues suffers from the significant shortcoming of a restriction to the analysis of access to managerial occupations (see Shephard and Levin, 1973; and Fernandez, 1975). But, it must be recognized that the exercise of authority in organizations is not

restricted to those who hold occupations that place them in the category of manager (Kluegel, 1975). Individuals who hold occupations other than manager are differentiated by their authority ranks. Furthermore, it should be recognized that all managers are not homogeneous in authority rank.

The only study to our knowledge that has examined authority differences among blacks and whites across the range of occupations is that of Wright and Perrone (1977). The focus of their study is, however, different from the focus of the research reported here. They examine black-white differences involving broad authority (class) categories, rather than authority rank. Since the main focus of their research was on the impact of these categories on income attainment, they offered no analysis of the sources of the black-white difference in authority position.

In this paper we present the results of our research on black-white differences in job authority. We begin by examining the magnitude of the gap in job authority between black and white males in our samples. We next examine the contribution of several determinants of authority attainment to the explanation of the black-white discrepancy in job authority. Finally, we turn to an analysis of the importance of racial differences in access to authority positions in accounting for the black-white income difference.

DATA

The analysis reported in this paper is based on data from two samples: a sample of 1,088 white nonfarm employed males (ages 20 to 65) from the state of Wisconsin, and a sample of 625 black, employed males from the Milwaukee area.¹ This research was conducted in the spring of 1973, and is part of a larger, nationwide study of social mobility (Featherman and Hauser, 1975).

¹ We limit our comparison to nonfarm white men to maximize the comparability of the two samples—since the sample of black men was drawn from an urban area, while the original white sample was drawn from both urban and rural areas.

At a minimum, we believe that the results of our analysis may be generalized to other northern industrial areas, whose industrial composition is like that of Wisconsin and the Milwaukee area. A recent comparison of racial stratification in the North and South (Hogan and Featherman, 1977), showing marked differences, cautions against making broad generalizations about black-white differences in job authority in the U.S. as a whole. Nevertheless, the results of our examination of black-white differences in aspects of socioeconomic achievement other than job authority (not reported here) do correspond closely to the results of national surveys (see Featherman and Hauser, 1976). This correspondence leads us to believe that although our results may not be viewed as strictly representative of the U.S. as a whole, they are not uncharacteristic of the general picture of black-white authority differences that would be found if a national study was to be undertaken. Measures of variables employed in our analysis are briefly described below.

Socioeconomic background. Respondents were asked to report on several characteristics of their parents, including father's occupation (at respondent's age 16), father's education, and family income (at age 16, age adjusted, in 1,000 dollar units). Information about father's occupation was used to create four additional variables: (1) a dummy variable (1,0) indicative of farm background; (2) a dummy variable for father's occupation manager; (3) a dummy variable for father's occupation foreman; and (4) a dummy variable for father self-employed.² These four variables were created to measure what may be called the "authority character" of father's occupation; this permitted us to determine if sons of fathers from such oc-

cupations experience any advantage in achievement within the authority ranks.

Respondent's education, occupation and work experience. Education is measured in total years. Information on respondent's current occupation was used to construct Duncan SEI scores. The most directly relevant measure from the standpoint of recruitment to authority positions may well be job tenure. That is, the amount of time that someone has held a specific job or has been in related jobs. Such a measure is unavailable in this study, and consequently our measure is based on our best available approximation to job tenure: total amount of time spent in the labor force.³ Work experience is measured in a metric of months, and equals the respondent's age less time spent in school and in military service.

Organizational characteristics. Information on both the size and industry category of the organization in which a respondent worked was obtained from his own report.⁴

Income. Individuals were asked to provide information on income from three

³ Since black men are more likely than white men to have discontinuous participation in the labor force, the lack of a measure of job tenure is troublesome. Some of the residual (unexplained) difference between blacks and whites in mean job authority (discussed subsequently) may be attributable to the discrepancy between our measure of work experience and actual job tenure.

⁴ Information on organization size was used for employees (non-self-employed) only. Respondents were asked to report on the size of the organization they worked for in terms of six categories of the number of people working for their employer. From these categories four dummy variables were formed for the following categories: (1) 1 to 24 employees, (2) 25 to 99 employees, (3) 100 to 499 employees, and (4) 500 or more employees. Ten industry categories were constructed using the census classification of major industry groups, and they are represented in subsequent regression equations by dummy variables. The categories are as follows:

1. mining and construction;
2. durable goods manufacturing;
3. nondurable goods manufacturing;
4. transportation, communication, and utilities;
5. wholesale trade;
6. retail trade;
7. business and repair service, personal service, entertainment and recreation services;
8. professional and related service;
9. public administration;
10. finance, insurance, and real estate.

² These variables were formed using census codes for detailed occupation categories. These four variables take on the following values:

- Farm Background = 1 if respondent's father held the occupation of farmer, and = 0 otherwise;
 Father Manager = 1 if respondent's father held the occupation of manager, and = 0 otherwise;
 Father Foreman = 1 if respondent's father held the occupation of foreman, and = 0 otherwise;
 Father Self-Employed = 1 if respondent's father was self-employed, and = 0 otherwise.

sources: (1) own business, professional practice or partnership; (2) income from wages and salaries; and (3) income received from their own farm. For the self-employed, income equals the total of all three sources. For the non-self-employed, income equals wages and salaries only.

Labor force participation. This measure refers to the amount of time spent in the labor force in 1972. Three variables were used to measure labor force participation: (1) number of weeks unemployed or on layoff from a job; (2) full time vs. part time work status; and (3) total weeks worked in 1972.

Union membership. A dummy (1,0) variable was formed indicating whether (1) or not (0) the respondent reported membership in a labor union.

Job authority. Our focus is on supervisory authority. A common approach to its measurement is based on position within the hierarchical chain of command as judged from the organization charts of specific organizations (see Tannenbaum et al., 1974). Unfortunately this approach has a significant drawback. Measuring job authority through position in the formal organization chart does not lend itself well to the mode of sampling individuals in studies of socioeconomic achievement.

We have employed a measure of job authority better suited to the sampling of individuals from a more broadly defined population. Our measure is based on respondents' replies to questions concerning whether or not they exercised various supervisory responsibilities as part of their jobs. These questions were used to form an additive index of job authority.⁵

⁵ The basis of the measure of job authority employed in this study is the respondent's report of whether or not he performs certain supervisory activities as part of his job. Specifically, respondents were asked the following set of questions:

1. I have authority to hire or fire others.
Yes (1) No (0)
2. I can influence or set the rate of pay received by others.
Yes (1) No (0)
3. Someone else influences or sets my rate or amount of pay.
Yes (0) No (1)
4. I supervise the work of others, that is, what they produce or how much.
Yes (1) No (0)

Since individuals at higher ranks in the authority hierarchy typically exercise more control over the major supervisory responsibilities in an organization (Dubin, 1958; Tannenbaum et al., 1974), the distribution of supervisory responsibilities itself will be hierarchically distributed. Thus the total number of supervisory responsibilities can be viewed as an indicator of hierarchical position. The index we have constructed takes on values from zero to five, with five representing the highest value. A brief discussion of the validity of this measure is presented in the appendix to this paper.⁶

An important general point about the measure of job authority we employ needs to be emphasized. It should be recognized that this measure does not tap all aspects of organizational authority. A useful distinction is between hierarchical and nonhierarchical authority (Thompson, 1961). Hierarchical authority refers to official organizational authority derived from the right to supervise and give orders to others. Nonhierarchical authority, on the other hand, refers to authority that arises unofficially from technical specialization or expertise. Our index measures organizational authority in the hierarchical sense only.

BLACK-WHITE DIFFERENCES IN AUTHORITY ATTAINMENT

As a starting point we present in Table 1 descriptive statistics for job authority by race. Because of the indication of change in the socioeconomic circumstances of younger black men we present separate information for two large age groups, as

5. Someone else supervises my work, that is, what I produce or how much.
Yes (0) No (1)

The scale of job authority used in this research was constructed by first assigning a value of one or zero to the "Yes" or "No" response to each question. A value of one indicates that an individual exercises greater control over the task-related behavior of others, while a value of zero indicates lesser control. The assigned values for the responses to each question are presented in parentheses next to each response. An individual's total score for job authority is formed by adding together these values.

⁶ An extended discussion of the validity of this measure is presented in Kluegel (1975).

Table 1. Values of the Mean and Standard Deviation (in Parentheses) for Job Authority, Authority Occupations and Other Socioeconomic Characteristics: Males in the Employed Civilian Labor Force by Race

	Black		
	Ages 20-44	Ages 45-65	Total
	(n=438)	(n=187)	(n=625)
Job Authority	.693 (1.001)	.622 (1.125)	.669 (1.032)
Manager *	.034 (.181)	.018 (.133)	.025 (.156)
Foreman	.043 (.202)	.011 (.104)	.034 (.181)
Self-Employed	.016 (.125)	.040 (.196)	.022 (.147)
Occupational Status	25.747 (18.341)	20.707 (15.041)	24.238 (17.514)
Income	7,205 (4,221)	8,103 (4,951)	7,486 (4,453)
Education	11.234 (12.209)	8.595 (3.332)	10.439 (2.861)

	White		
	Ages 20-44	Ages 45-65	Total
	(n=712)	(n=376)	(n=1,088)
Job Authority	1.216 (1.421)	1.549 (1.597)	1.335 (1.494)
Manager *	.105 (.307)	.114 (.318)	.109 (.312)
Foreman	.036 (.186)	.052 (.222)	.042 (.200)
Self-Employed	.087 (.282)	.149 (.356)	.109 (.312)
Occupational Status	40.698 (24.534)	42.965 (23.787)	41.526 (24.305)
Income	10,166 (7,034)	12,223 (8,174)	10,783 (7,459)
Education	12.731 (2.260)	11.472 (3.011)	12.264 (2.626)

* Proportion of non-self-employed managers only.

well as for the total samples of both races. The entries in Table 1 reveal a marked discrepancy in job authority. For the total samples the job authority of black men averages roughly half that of white men. Also arrayed in Table 1 are the proportions by race of men who hold what may be called "authority occupations"—that is, the occupations of manager, foreman, and the status of being self-employed.⁷ Again, as in the case of mean

job authority, a marked discrepancy in favor of whites is shown.

The mean values of occupational status, income, and education (Table 1) show in this sample an improvement in the general socioeconomic circumstances of younger black men that parallels the improvement found in national studies. Our data is indicative of a similar improvement among younger black men in job authority.⁸ With the exception of self-employed status, the same conclusion applies to the proportion of younger black men found in authority occupations.

In Table 2 we present partial regression

⁷ We do not intend to imply by the use of the term "authority occupation" that there is a meaningful conceptual distinction between authority occupations and job authority as a dimension of position within the social division of labor. We present information on authority occupations principally because of the history of interest in these distinct occupations in prior discussion of black-white differences in socioeconomic achievement. We are in general accord with the sentiment expressed by one reviewer that the task before us, if we are to develop a theory of income determination, is the identification of dimensions of position within the social division of labor. We believe that job authority is one such dimension and one goal of this paper is to present research that we hope will illustrate the importance of giving serious attention to this dimension in the development of a theory of income determination.

⁸ Since authority attainment takes place over an extended period in the socioeconomic life cycle, one cannot assume that the size of the black-white gap in job authority will remain the same when those individuals in the 20- to 44-year-old age group grow older. This realization stands as a cautionary note to the interpretation of improvement in the job authority status of younger black men. Nevertheless, we believe that such an inference is still sound, since although among white men those in the older age group have a higher average value of job authority than the younger group, among black men the opposite order of average job authority by age group holds.

coefficients for equations relating variables measured in this study that may be thought of as potential determinants of authority position to our measure of job authority. The factors represented in these equations generally have received at least speculative attention in prior discussion of the possible sources of black-white differences in job authority. Equations were estimated both within the total sample and within the younger age group (20-44) for each race. The furthest right hand column (equation III) presents values of partial regression coefficients, estimated within the total samples, for the determinants of job authority excluding measures of occupational position.

The principal distinction between the three equations is that in the third equation we are examining the direct effects of the listed factors on the attainment of jobs characterized by the exercise of authority; while in the first two equations we are examining the impact of factors on within-occupation variation in job authority. The estimation of equation III permits us to treat an area of conceptual overlap between our study and prior research that has examined racial differences in the attainment of occupational status. There is a moderate correlation between job authority and occupational status (among white men it equals .414), reflecting the fact that authority is one of the dimensions underlying occupational status. But, there are other factors that contribute to occupational status including the technical expertise required to perform an occupation and the degree to which the product produced by members of the occupation is valued by society. Consequently, prior research does not permit an explicit assessment of the difference in the type of jobs held by blacks and whites uniquely attributable to exclusion from authority positions. Equation III can be used for this assessment.

There is also an aspect of racial differences in job authority that is not tapped by studies of occupational status. Such studies do not deal with black-white differences in job authority among incumbents of similar occupational positions. To examine this aspect of job authority differences, we include in equations I and II

four measures of occupational position: occupational status, and three dummy (1,0) variables indicating whether one holds a managerial or nonmanagerial occupation, whether one is a foreman or not, and whether one is self-employed or not self-employed.⁹ Occupational status score provides a rough means of controlling for aspects of occupational position, such as technical expertise, that affect access to authority positions.¹⁰ The three dummy variables are included because of our judgment that occupational status does not exhaustively represent the influence of aspects of occupational position represented by these variables.

From Table 2 we can make the following initial observations about sources of the job authority gap. First, apart from its influence through education, racial differences in socioeconomic background have virtually no effect on the job authority discrepancy. Socioeconomic background has no significant direct effect among blacks or whites on either access to jobs characterized by the exercise of authority or on within-occupation variation in job authority. Second, equation III indicates that there is a clear effect of organizational characteristics on the attainment of jobs characterized by the exercise of authority. The relevant coefficients in this equation represent deviations of the mean job authority within each size and industry category from the grand mean of job authority, with differences among size and industry categories in the socioeconomic background, education and work experience of their workforces controlled

⁹ These three dummy variables were formed using census codes for detailed occupation categories, as follows:

1. manager = 1, nonmanager = 0;
2. foreman = 1, nonforeman = 0;
3. self-employed = 1, non-self-employed = 0.

¹⁰ We emphasize that our intent here is to control for aspects of occupational position that on the average, across individual incumbents of occupations, affect access to authority positions. Thus when we speak of the impact of the technical expertise of occupations on access to authority we refer not to a property of an individual, but of the occupation that individual holds. Holding the occupation of lawyer, for example, provides heightened access to authority positions for all incumbents by virtue of the role of legal knowledge in decision making in organizations.

Table 2. Standardized and Metric (in Parentheses) Partial Regression Coefficients from Equations for the Determinants of Job Authority: Men in the Experienced Civilian Labor Force By Race

Independent Variables	White		
	I (Total)	II (20-44)	III
FAM INC	.014 (.002)	-.001 (-.001)	.035 (.005)
F's OCC	-.001 (-.001)	.072* (.004)	.021 (.001)
F's OCC MAN	.023 (.096)	.002 (.007)	.040 (.164)
F's OCC FORE	.019 (.140)	.004 (.030)	-.002 (-.019)
F's OCC SELF	.029 (.093)	.067 (.210)	.078 (.248)
F's EDUC	.018 (.007)	-.045 (-.018)	.027 (.011)
FARM BACK	-.008 (-.034)	-.022 (-.090)	-.009 (-.033)
R's EDUC	-.006 (-.003)	.040 (.025)	.173* (.098)
WORK EXP	.068* (.007)	.106* (.019)	.234* (.024)
R's OCC	.245* (.015)	.210* (.012)	
R's OCC MAN	.202* (.818)	.236* (.960)	
R's OCC FORE	.127* (.944)	.122* (.925)	
R's OCC SELF	.497* (2.377)	.466* (2.350)	
UNION	-.099* (-.304)	-.089* (-.261)	-.257* (-.792)
ORG SIZE 1 (251) *	(.035)	(.089)	(-.320)
ORG SIZE 2 (210)	(.204)	(.160)	(.265)
ORG SIZE 3 (171)	(-.136)	(-.148)	(-.044)
ORG SIZE 4 (338)	(-.084)	(-.091)	(.094)
MAN DUR (278)	(.032)	(.011)	(-.225)
MAN NONDUR (136)	(-.080)	(-.115)	(-.485)
TRAN COMM UTS (105)	(-.108)	(-.119)	(-.307)
WHOLE (50)	(-.209)	(.135)	(-.081)
RETAIL (115)	(.050)	(.054)	(.702)
BUS PERS SERV (47)	(.156)	(.058)	(.907)
PROF SERV (46)	(-.080)	(-.195)	(.126)
PUB ADMIN (133)	(.069)	(.014)	(-.107)
FIN REAL (65)	(-.070)	(-.033)	(.128)
MIN CON (111)	(.092)	(.178)	(.462)
R ² (constant)	.598 (.167)	.596 (-.033)	.227 (.140)

for.¹¹ Although there is some effect of organizational size category among whites, in general it appears that industry category has the strongest impact on the overall chance of obtaining authority positions. The number of workers within each

size and industry are presented in parentheses next to the variable names in Table 2. They reveal that blacks in this sample are found disproportionately in comparison to whites in industries that offer the poorest relative chances for authority attainment (principally in manufacturing industries).

¹¹ The regression coefficients for the organizational size and industry categories reported in Table 2 were estimated by deleting one category from each set of categories (via dummy variable coding) and then transformed to obtain the deviation of the adjusted mean of job authority within each category from the overall (grand mean) of job authority (Malichar, 1965). F-tests of the increment to explained variance show that for whites both organizational size and industry categories significantly affect job authority ($p < .01$) in equation III. For blacks, however, only industry categories significantly affect job authority. In equations I and II neither organizational size nor industry categories have significant direct effects on job authority for whites or blacks.

From equations I and II and from information taken from Table 1, we can conclude that components of the difference in job authority are contributed by both racial differences in occupational position and by differences in authority attainment within occupations. The former component reflects the reported differences (Table 1) in mean occupational status and in the proportion manager, foreman, and self-employed. The latter component is demonstrated by the racial difference in the effect of occupational

Independent Variables	Black		
	I (Total)	II (20-44)	III
FAM INC	-.009 (-.001)	-.035 (-.006)	.020 (.003)
F's OCC	.035 (.003)	.003 (.001)	.082 (.007)
F's OCC MAN	-.030 (-.178)	-.002 (-.010)	-.053 (-.314)
F's OCC FORE	.064 (1.095)	.076 (1.062)	.078 (1.336)
F's OCC SELF	.034 (.084)	.028 (.067)	-.018 (-.044)
F's EDUC	.052 (.014)	.051 (.015)	.011 (.003)
FARM BACK	.012 (.028)	-.005 (-.011)	.036 (.085)
R's EDUC	.102* (.037)	.067 (.031)	.309* (.112)
WORK EXP	.088* (.007)	.097 (.012)	.221* (.018)
R's OCC	.118* (.007)	.110* (.006)	
R's OCC MAN	.182* (.974)	.215* (1.180)	
R's OCC FORE	.223* (1.273)	.257* (1.261)	
R's OCC SELF	.337* (2.346)	.371* (2.962)	
UNION	-.098* (-.202)	-.100* (-.201)	-.147* (-.303)
ORG SIZE 1 (104) *	(.234)	(.294)	(.162)
ORG SIZE 2 (89)	(.047)	(.009)	(.026)
ORG SIZE 3 (110)	(-.011)	(-.023)	(.009)
ORG SIZE 4 (304)	(-.089)	(-.094)	(-.065)
MAN DUR (263)	(-.067)	(-.069)	(-.156)
MAN NONDUR (83)	(-.026)	(-.022)	(.006)
TRAN COMM UTS (44)	(.016)	(-.090)	(-.099)
WHOLE (24)	(.446)	(.319)	(.319)
RETAIL (53)	(.048)	(.035)	(.541)
BUS PERS SERV (15)	(.345)	(.393)	(.547)
PROF SERV (49)	(.054)	(.033)	(.248)
PUB ADMIN (31)	(-.116)	(.011)	(-.136)
FIN REAL ^b (5)
MIN CON (40)	(.183)	(.241)	(.186)
R ² (constant)	.410 (-.382)	.444 (-.271)	.153 (-1.079)

* Category n.

^b Because of the small number of cases, individuals in this category were excluded.

* = $p < .05$.

Note: FAM INC=family income at age 16 in age adjusted dollars, \$1,000 units; F's OCC=father's occupational status (Duncan SEI scores); F's OCC MAN=father's occupation manager; F's OCC FORE=father's occupation foreman; F's OCC SELF=father self-employed; F's EDUC=father's education in years, 17 or more=17; FARM BACK=farm background; R's EDUC=respondent's education; WORK EXP=work experience (measured in months, multiplied by 12 for presentation in this table); R's OCC=respondent's occupational status (Duncan SEI scores); R's OCC MAN=respondent's occupation manager; R's OCC FORE=respondent's occupation foreman; R's OCC SELF=respondent self-employed; ORG SIZE 1=1 to 24 employees; ORG SIZE 2=25 to 99 employees; ORG SIZE 3=100 to 499 employees; ORG SIZE 4=500 or more employees; MAN DUR=durable goods manufacturing; MAN NONDUR=non-durable goods manufacturing; TRAN COMM UTS=transportation, communication, and utilities; WHOLE=wholesale trade; RETAIL=retail trade; BUS PERS SERV=business and repair service, personal service, entertainment, and recreation services; PROF SERV=professional and related services; PUB ADMIN=public administration; FIN REAL=finance, insurance and real estate; MIN CON=mining and construction.

status on job authority.¹² The regression coefficient for the influence of occupational status on job authority is roughly

half as large in both the total sample and in the younger age group of black men, as it is in the parallel samples of white men.

¹² Tests of the difference between regression coefficients show that, for both the total sample ($t = 2.250$, $p < .05$) and for the 20-44 age group ($t = 1.731$, $p < .10$), the effect of respondent's occupational status on job authority differs significantly between blacks and whites.

Finally, we have included union membership as a possible determinant of authority attainment, largely because of its possible relationship to perceived loyalty to the organization. Given the tendency for union membership to be a characteris-

tic of incumbents of lower-status occupations, the impact of union membership as an indicator of perceived loyalty is best assessed by its effect on authority attainment, with occupational position controlled for. From equations I and II we find that union membership, consistent with the hypothesis that it is viewed by employers as an indicator of disloyalty, has a negative effect on job authority. The black sample shows a higher proportion of union members (.495) than does the white sample (.380), suggesting that this difference may contribute to the explanation of the authority gap.

As a means of assessing the relative importance of each source discussed above we have employed the technique of indirect standardization (Duncan, 1969; Winsborough and Dickinson, 1971; Althauser and Wigler, 1972). The results of the application of this technique are arrayed in Table 3. The entries in this table were generated by substituting black-white mean differences into equations for the process of authority attainment estimated within the black sample.¹³ We assume that the determinants of job authority examined in this analysis stand in the following causal order (from most prior to most subsequent): socioeconomic background, education, organizational characteristics, occupational position, union membership, and job authority. The process of substitution was performed cumulatively. Thus the entries in Table 3 represent the total effect—both direct and indirect through causally subsequent factors—of the factors listed in the left-hand margin of this table. Because we have found no direct effect of socioeconomic background on job authority, and for simplicity of computation, we allow the indirect effects of socioeconomic background to be absorbed in causally

Table 3. Components of the Black-White Mean Difference in Job Authority

Component	Value	% of Difference
Organizational		
Characteristics	.106	15.9
Education	.192	28.8
Union Membership	.027	4.1
Occupational Position	.210	31.5
Within-Occupation		
Exclusion	.131	19.7
Total	.666	100.0

subsequent factors. The number of determinants of job authority considered and the complexity of their interrelationship make the entries in Table 3 approximate estimates. Nevertheless, they do provide rudimentary information about the relative importance of sources of the mean job authority gap between white and black men.

As in the case of racial differences in other aspects of socioeconomic achievement, the black-white difference in job authority is only moderately determined (approximately 30% of the mean difference) by differences in human capital (as represented by years of education). Depending upon the assumptions made, the remaining 70% of the job authority discrepancy can be attributed, all or in part, to racial discrimination.

In this regard we point out that the component due to occupational position represents the effect of what may be thought of as "occupational discrimination" (Duncan, 1969). That is, the cumulative procedure used absorbs occupational differences due to educational differences in the education factor. Furthermore, a substantial part of the occupational position component itself is attributable to exclusion from authority occupations.¹⁴ We can specify that one aspect of what frequently has been given the rather undifferentiated label of "occupational discrimination" is the exclusion of blacks

¹³ The equations estimated within the black sample were used because of the reported lesser effect of occupational status on job authority among black men. We also estimated the components in Table 3 by substituting black-white mean differences into equations estimated within the white samples. Corresponding estimates of the components of the job authority gap due to education, organizational characteristics, and union membership from both methods are approximately equal.

¹⁴ Indeed it may well be that a major source of exclusion from authority occupations is exclusion from promotions that involve a change of occupations (e.g., intragenerational mobility into the occupation of manager).

from occupations that are principally characterized by the exercise of job authority. Also, differences in the characteristics of organizations that employ blacks and whites are rooted in part in the history of race relations in this country (Bonacich, 1976). Thus one can construct a reasonable argument for considering the effect of these two factors as aspects of racial discrimination in job authority.

The remaining components present greater problems of classification. The component we have labeled "within-occupation exclusion from authority" must ultimately be recognized as a residual factor. Although we have found that black men at the same level of occupational status experience poorer access

to authority positions than white men, explicit racial discrimination is only one factor that may account for within-occupation differences. We do know, however, that education, socioeconomic background, and the gross organizational characteristics we have measured in this study do not account for the within-occupation difference in job authority. Finally, we admit of knowing no conventional wisdom on the legitimacy of union membership as a criteria for exclusion from authority positions.

THE COST OF THE JOB AUTHORITY GAP

In this section we address the question of how much of the total observed income

Table 4. Standardized and Metric (in Parentheses) Partial Regression Coefficients from Equations for the Determinants of Income: Men in the Employed Civilian Labor Force by Race

Independent Variables	White		
	IV (Total)	V (20-44)	VI
R's EDUC	.112* (3.169)	.038 (1.186)	.246* (6.977)
WORK EXP	.083* (.449)	.121* (1.080)	.140* (.753)
UNION	.103* (15.876)	.114* (16.627)	.140* (21.501)
JOB AUTH	-.150* (-7.485)	-.170* (-8.430)	.329* (16.429)
R's OCC MAN	-.060 (-12.153)	-.071 (-14.275)	
R's OCC FORE	-.048 (-17.929)	-.041 (-15.340)	
R's OCC SELF	.032 (7.542)	.099* (24.775)	
R's OCC	.060 (.183)	.136* (.390)	
UNEMP	-.023 (-.200)	-.012 (-.161)	
FULL PART	.111* (38.461)	.135* (38.289)	
TOTAL WEEKS	.243* (1.462)	.209* (1.178)	
R's OCC X AUTH	.495* (.382)	.410* (.317)	
R ² (constant)	.363 (-38.190)	.341 (-13.845)	.229 (-2.401)
Independent Variables	Black		
	IV (Total)	V (20-44)	VI
R's EDUC	.096* (1.507)	.129* (2.463)	.218* (3.393)
WORK EXP	.173* (.610)	.207* (1.104)	.217* (.756)
UNION	.219* (19.497)	.223* (18.845)	.351* (31.774)
JOB AUTH	.003 (.123)	-.026 (-1.114)	.151* (6.507)
R's OCC MAN	.017 (3.945)	-.054 (-12.422)	
R's OCC FORE	-.057 (-14.056)	-.053 (-11.046)	
R's OCC SELF	-.129* (-38.783)	-.173* (-58.319)	
R's OCC	.136* (.345)	.034 (.079)	
UNEMP	.052 (.205)	.093* (.333)	
FULL PART	.056* (10.653)	.021 (3.612)	
TOTAL WEEKS	.588* (1.538)	.612* (1.506)	
R's OCC X AUTH	.127* (.114)	.219* (.176)	
R ² (constant)	.586 (-43.164)	.599 (-48.042)	.254 (9.012)

* = $p < .05$.

Note: JOB AUTH=job authority; UNEMP=weeks unemployed or on layoff from a job; FULL PART=full vs. part-time work status; TOTAL WEEKS=total number of weeks worked in 1972; R's OCC X AUTH=occupational status times job authority. For a definition of remaining variables see Table 2.

gap among black and white men in this data is attributable to the exclusion of black men from authority positions. Table 4 contains values of partial regression coefficients from equations for the determinants of income. In all the reported equations organizational characteristics were controlled for, but for simplicity of presentation the regression coefficients for organizational characteristics are not presented in Table 4.

We can make a parallel distinction to that made in the analysis of authority attainment between the total effect of the exercise of authority on income attainment and the effect of within-occupation differences in job authority. To facilitate the examination of these different aspects of the cost of exclusion from authority we estimated equation VI for both races. Equation VI differs from equations IV and V in three ways. First, measures of occupational position are not included in the equation. This exclusion permits us to assess the total effect of holding jobs characterized by the exercise of authority on income attainment. Second, equations IV and V also include measures of labor force participation, excluded from equation VI. Third, we have included in equations IV and V a linear interaction term (ROCCAUTH) to express, as we reported in prior research (Kluegel, 1975), the fact that a unit increase in job authority is worth more income at higher occupational status levels. In short, equation VI represents the total effect of holding jobs characterized by the exercise of authority—both direct and indirect effects through labor force participation—averaged across levels of occupational status.

From Table 4 we can make the following observations. First, it is clear that the exercise of authority is an important determinant of income. From equation VI we can see, however, that black men receive far less than white men for equal increases in job authority, on the average across occupations.¹⁵

Equations IV and V demonstrate the previously noted characteristic that income returns to unit increases in job au-

Table 5. Partial Regression Coefficients for the Effect of Job Authority on Income within Selected Occupational Status Levels by Race

Occupational Status	Black	White
80	9.24*	23.07
60	6.96	15.43
40	4.68	7.80
30	3.54	3.97

* Income is measured in 100 dollar units.

thority are greater at higher levels of occupational status. The same pattern of relationship is demonstrated for blacks and whites, with the important exception that the slope for whites is substantially steeper than the slope of the job authority-times-occupational status term for blacks.¹⁶ Thus the discrepancy in black-white income returns is greater at higher than at lower occupational status levels. To illustrate the nature of this discrepancy we present in Table 5 for the black and white total samples values for the partial regression coefficient of income on job authority within selected occupational status levels.¹⁷

A final observation made from Table 4 is that younger black men show a smaller discrepancy in income returns to job authority in a comparison to the younger group of white men than the discrepancy found in a comparison of total samples. This finding may be seen as consistent with the recent evidence of a general improvement in the socioeconomic circumstances of blacks.

¹⁶ Values of t-tests for the difference in these regression coefficients equal 3.167 ($p < .01$) and 1.726 ($p < .10$), respectively, for the total group and age 20-44 group.

¹⁷ To obtain the values arrayed in Table 5, values of the specified status score were multiplied by the regression coefficients for the occupational status-times-job authority (ROCCAUTH) term estimated within the black and white samples. To the resultant quantities we then added the regression coefficients for the respective additive effects of job authority (JOB AUTH) estimated within each sample. Thus, for example, the regression coefficient of income on job authority for black men with a status score of 80 equals,

$$80 \times .114 + .123 = 9.24.$$

Since income has been coded in units of 100 dollars, we thus find that a one-unit increase in job authority among black men in occupations with a status score of 80 is worth 924 dollars.

¹⁵ The value of t in the test of the difference of regression coefficients equals 3.905, $p < .01$.

To provide a summary assessment of the impact of the exclusion from job authority on the black-white difference in mean income we again employ indirect standardization. To generate the entries in Table 6, values for black-white mean differences were substituted into equations for the determinants of income from the total sample of whites.¹⁸ From Table 6 we see that the exclusion of black men from job authority on the average costs black men roughly 1,100 dollars per year. This figure was obtained by using equation VI and hence includes a component due to the exclusion from occupations that provide access to authority positions, as well as a component for the cost of exclusion from job authority within occupations. Estimates of these subcomponents are presented in Table 6 under the respective labels of "between occupations" and "within occupations." The total cost of exclusion from job authority is only slightly less than the total (both direct and indirect through occupational position) cost of black-white differences in years of education. The cost of exclusion from authority is greater than our estimate of the cost of another important source of the black-white income difference, the difference in labor force participation.

DISCUSSION

The results of our analysis, coupled with prior findings, enable us to offer the following general insights concerning the

Table 6. Components of the Black-White Mean Difference in Income

Component	Value	% of Difference
Total Job Authority	1,094	33.2
a. Between Occupation	674	20.4
b. Within Occupation	420	12.8
Education	1,273	38.6
Labor Force Participation	868	26.3

than job authority. Prior research has shown substantial movement of blacks into what may be viewed as better jobs within the blue-collar sector. Between 1962 and 1973 a large increase was realized in the percentage of black men who hold jobs within the broad occupational category of operatives and transport workers (Hauser and Featherman, 1974:252-3). Such jobs are often found in large, unionized firms, where workers realize the benefits of membership in industrial unions (Ashenfelter, 1972) and the further income benefits that accompany employment by large firms. However, these jobs present little opportunity for advancement within authority hierarchies. But, even when blacks come to occupy jobs that usually present better opportunities for advancement our research shows that their access to job authority is less than that of whites. Briefly stated, a major aspect of the current racial difference in socioeconomic achievement is exclusion from authority, both in the recruitment to authority occupations and

persistent difference between blacks and whites in socioeconomic achievement. First, it appears that the principal route of progress toward equality of black-white socioeconomic achievement has been through improvements in aspects of the nature of jobs held by black men other

in an important source of within-occupation achievement.

This exclusion from authority is costly, accounting for approximately one-third of the total black-white income gap in our data. Furthermore, exclusion from authority is more costly at higher levels of occupational status. This finding, combined with the finding of a difference between the white and black samples in the effect of occupational status on job authority may be of value in helping to explain a characteristic of the black-white income difference reported in two prior

¹⁸ The equations for the white sample were used because we believe they provide our best estimate of the total cost of exclusion from authority. The total cost can be subdivided into two components: a component reflecting the difference between blacks and whites in income returns to job authority, and a component reflecting the black-white difference in

higher educational levels. On the basis of our results we speculate that this finding can in part be attributed to the higher cost of exclusion from authority within occupations of the status levels that typically accompany higher educational levels.

We have some evidence of recent improvement in the job authority status of black men. Black men in the younger age group are found more frequently than those in the older age group in the occupations of manager and foreman, and have higher average job authority. Furthermore, the noted difference in income returns for increases in job authority is smaller in the younger group. There are aspects of our findings, however, that must be seen as qualifying any conclusion of improvement. First, there is still a substantial difference between the whites and blacks in the younger age group in both average job authority and in income returns to job authority. Second, the same pattern (as indicated by the partial regression coefficient for the effect of occupational status on job authority) of within-occupation restriction in access to authority among black men exists within the total sample and the younger age group of black men. Finally, our data indicates no change (Table 1) in the self-employment status of blacks.

Any conclusion about trends in the authority status of black men on the basis of this data must, of course, be drawn with some caution since they are based on cross-sectional comparisons. There is an additional restriction on making inferences about likely trends in authority attainment among black men that lies in a lack of a comprehensive understanding of the process of achievement within the authority hierarchies of organizations. As we have argued elsewhere (Kluegel, 1975) there is a paucity of systematic research on this topic. The results of our research do, however, provide some initial insights into this process.

We have seen that what may be called "structural factors" have a substantial impact on access to authority positions. The characteristics of the employing organization and the occupational position held are important determinants of one's chance of attaining job authority. These

factors can be considered to be structural determinants because they reflect the influence of the social organization of work. The technical expertise of certain occupations is one aspect of the relationship of occupational position to the social organization of work. Such expertise can place incumbents of an occupation in an advantaged position to make decisions in a specialty area and increase access to authority positions. And, both the type of technology employed (roughly indicated by industry group in our analysis) by an organization and its size influence the supply of authority positions available (Hall, 1972). Whatever social factors influence the allocation of workers to occupational positions and to different types of organizations thus will be consequential for their attainment of job authority. Race is such a factor.

A potentially important area in the study of the process of authority attainment that we could not address directly in this research is that of the influence of beliefs, values, and attitudes. It has been argued that political beliefs and values (Dalton, 1959; Grusky, 1965), and such perceived attitudes as loyalty and commitment to the organization's goals and values (Edwards, 1975) are determinants of promotion within authority hierarchies. Some of the results of our research do permit us to speak indirectly to the nature of the influence of these unmeasured factors. Of particular interest in this regard is the influence of education and socioeconomic background on authority attainment.

Our analysis shows that the impact of education on job authority is essentially indirect. Increases in years of education present access to occupations that in turn provide better chances for acquiring authority positions. But, as can be seen from Table 3, there is a negligible effect of years of education on within-occupation variation in job authority. Furthermore, we find no direct effect of socioeconomic background on job authority. These findings suggest that if beliefs and values play a role in job authority attainment, then they are beliefs, values, and attitudes that are not associated with educational level or socioeconomic background. If such

attributes were acquired through formal education or socioeconomic background experiences, then one might well expect to find within-occupation effects of these variables. We can further extend this argument to consider the hypothesis that differences in beliefs, values and attitudes account for a significant part of the unexplained black-white difference in job authority. If indeed such an hypothesis was true, then, whatever their nature, these beliefs, values and attitudes would appear to have their origin, at least in part, in experiences associated with race and not in experiences associated with either formal education or socioeconomic background.

CONCLUSION

Our intent in this paper has been to provide what we consider to be necessary first steps toward understanding the role of job authority as a source of the persistent black-white difference in socioeconomic achievement. As a result the research reported here has been of a somewhat exploratory nature. Refinements of research design and measurement, such as the use of longitudinal designs and a more direct measure of relevant work experience, are called for in future research. In conclusion, we offer suggestions for what we believe to be some important next steps in research on this topic.

In general, our suggestions relate to a perceived need for further studies of the dynamics of racial exclusion from authority. In this regard, we need to address such questions as whether racial exclusion is rooted in a clash between the beliefs, attitudes and values of black Americans and the basic value premises of capitalistic firms; or in straightforward racial prejudice operationalized in organizations by such rationalizations as one cannot place black men in positions of authority over whites. In this vein one must also consider the possibility that there are skill areas, unmeasured by years of education, that are important for effective functioning in authority positions. The question of whether black men have received adequate training in such skills (and indeed the question of how such training is ac-

quired) is of special importance in light of the controversy surrounding affirmative action programs (Glazer, 1975).

Our finding that black men receive substantially less of an income return for increases in job authority presents an additional question for further research. Prior research offers some potential explanations that should be examined. First, the difference in income return may lie in the unique nature of the entrepreneurial activities available to black men (Light, 1972; Aldrich and Reiss, 1976). Second, it may have its source for employees of organizations in the allocation of blacks to internal niches in complex organizational hierarchies that present low income returns to job authority for any occupant (Edwards, 1975). As a third explanation, there is, of course, the possibility of strictly discriminatory pay differentials.

APPENDIX

Although we believe that the premise upon which we have based our measure of job authority is sound, it remains to be demonstrated empirically that it, in fact, yields a valid measure of this construct. One simple method would be to gauge its predictive validity by asking the question, "Does this scale accurately predict an individual's position in the formal organizational chart of supervisory authority?" Since we lack the necessary data to assess predictive validity we rely instead on the notion of construct validation, and ask, "Does this scale behave as if it measures job authority?" To answer this question we employ the following five criteria:

1. For individuals who are employees of organizations a measure of job authority should have a pyramidal distribution. One should find few individuals occupying the highest scale value, with progressively more individuals at each successive scale value until we come to the lowest and most populated scale value.
2. The second criterion is given by research on the structure of organizations. A common finding is that the size of an organization is inversely related to the proportion of authority positions (Hall, 1972)—as the size of an organization increases the proportion of authority positions (to nonauthority positions) decreases. Thus one should expect to find that a measure of job authority yields the same inverse relationship with organizational size.
3. A valid measure of job authority should display an increasing monotonic relation to income.
4. A measure of authority rank should distinguish between respondents who are self-employed and respondents who are not.

Table A. Data for the Assessment of the Validity of the Job Authority Measure

Occupation	Job Authority		
	Mean	Standard Deviation	N
Self-Employed Professionals	4.000	1.219	29
Salaried Professionals	1.356	1.137	177
Salaried Managers	2.459	1.164	111
Self-Employed Managers	4.254	1.293	64
Sales	1.433	1.249	67
Clerical	.834	1.124	103
Crafts	1.104	1.154	359
Operatives	.444	.859	557
Service	.732	.970	106
Laborers	.462	.778	158

Scale Values of Job Authority	Frequency	Mean Income
5	7	\$20,240
4	32	12,451
3	115	14,300
2	207	10,392
1	419	9,167
0	801	8,029

Organization Size	Mean Job Authority	N
1 to 24 Employees	1.029	376
25 to 99 Employees	1.100	300
100 to 499 Employees	.720	282
500+ Employees	.671	578

Class of Worker	Mean Job Authority	N
Self-Employed	4.038	129
Non-Self-Employed	.844	1581
1. Private	.816	1332
2. Government	1.005	249

5. A valid measure of job authority should confirm our expectations about the distribution of authority among occupational groups. It should, for example, distinguish between white-collar and blue-collar occupations.

Data for the assessment of how well these criteria are met is presented in Table A. Looking first at the means and standard deviations for twelve major occupational groups we find that the mean values correspond closely to the ranking of occupations by job authority that we might expect. Note as well, as indicated by the standard deviation of job authority within each group, that considerable variance in job authority lies within each group. Values of this scale distinguish among individuals within each occupational group as well as between occupations. Turning to the remaining four criteria we find that the measure of job authority shows the expected inverse relationship with organizational size and that it clearly distinguishes between the self-employed and the non-self-employed. The frequency distribution for the scale values demonstrates the expected pyramidal distribution among employees of organizations. Finally, with the exception of the scale value of three where there is a deviation from the overall pattern, income displays an increasing monotonic relationship to values of job authority.

In summary, this measure of job authority appears to meet quite adequately four of the five criteria, and only somewhat less adequately to meet the fifth. On the balance, we believe that the data presented demonstrate the overall validity of our measure of job authority and justify its use in this research.

REFERENCES

- Aldrich, H., and A. J. Reiss, Jr.
 1976 "Continuities in the study of ecological succession: changes in the race composition of neighborhoods and their businesses." *American Journal of Sociology* 81:846-66.
- Althausen, R. P., and M. Wigler
 1972 "Standardization and component analysis." *Sociological Methods and Research* 1:97-135.
- Ashenfelter, O.
 1972 "Racial discrimination and trade unions." *Journal of Political Economy* 80:462-3.
- Bonacich, E.
 1976 "Advanced capitalism and black/white relations in the U.S.: a split labor market in-

- terpretation." *American Sociological Review* 41:34-51.
- Collins, R.
1971 "Functional and conflict theories of educational stratification." *American Sociological Review* 36:1002-19.
- Dalton, M.
1959 *Men Who Manage*. New York: Wiley.
- Dubin, R.
1958 *The World of Work*. Englewood Cliffs: Prentice-Hall.
- Duncan, O. D.
1969 "Inheritance of poverty or inheritance of race?" Pp. 85-110 in D. P. Moynihan (ed.), *On Understanding Poverty*. New York: Basic Books.
- Edwards, R. C.
1975 "The social relations of production in the firm and labor market structure." Pp. 3-26 in R. C. Edwards, M. Reich, and D. M. Gordon (eds.), *Labor Market Segmentation*. Lexington: Heath.
- Featherman, D. L., and Robert M. Hauser
1975 "Design for a replicate study of social mobility in the United States." Pp. 219-51 in K. C. Land and S. Spilerman (eds.), *Social Indicator Models*. New York: Russell Sage.
1976 "Changes in the socioeconomic stratification of the races, 1962-73." *American Journal of Sociology* 82:621-51.
- Ferman, L. A.
1968 *The Negro and Equal Employment Opportunities*. New York: Praeger.
- Fernandez, J. P.
1975 *Black Managers in White Corporations*. New York: Wiley.
- Glazer, N.
1975 *Affirmative Discrimination*. New York: Basic Books.
- Gordon, D. M.
1972 *Theories of Poverty and Underemployment*. Lexington: Heath.
- Grusky, O.
1965 "Career mobility and managerial political behavior." *Pacific Sociological Review* 8:82-9.
- Hall, R. H.
1972 *Organizations*. Englewood Cliffs: Prentice-Hall.
- Hauser, R. M., and D. L. Featherman
1974 "White-nonwhite differentials in occupational mobility among men in the United States, 1962-1972." *Demography* 11:247-65.
- Haworth, J. G., J. Gwartney, and C. Haworth
1975 "Earnings, productivity, and changes in employment discrimination during the 1960's." *American Economic Review* 68:158-68.
- Hogan, D. P., and D. L. Featherman
1977 "Racial stratification and socioeconomic change in the American North and South." *American Journal of Sociology* 83:100-26.
- Jencks, C., M. Smith, H. Acland, M. Bane, D. Cohen, H. Gintis, B. Heyns, and S. Michelson
1972 *Inequality*. New York: Basic Books.
- Johnson, M. P., and R. R. Sell
1976 "The cost of being black: a 1970 update." *American Journal of Sociology* 82:183-90.
- Kanter, R. M.
1975 "Women and the structure of organizations: explorations in theory and behavior." Pp. 35-74 in M. Millman and R. M. Kanter (eds.), *Another Voice*. Garden City: Anchor Books.
- Kluegel, J. R.
1975 *Job Authority and Social Inequality*. Ph.D. dissertation, Department of Sociology, University of Wisconsin, Madison.
- Light, I. H.
1972 *Ethnic Enterprise in America*. Berkeley: University of California Press.
- Malichar, E.
1965 "Least-squares analysis of economic survey data." *Proceedings of the Business and Economic Statistics Section of the American Statistical Association*, Philadelphia.
- Masters, S. H.
1975 *Black-White Income Differentials*. New York: Academic Press.
- Morgan, J. S., and R. J. Van Dyke
1970 *White-Collar Blacks*. American Management Association.
- Rees, A., and G. Shultz
1970 *Workers and Wages in an Urban Labor Market*. Chicago: University of Chicago Press.
- Shephard, W. G., and S. G. Levin
1973 "Managerial discrimination in large firms." *Review of Economic Statistics* 55:412-22.
- Siegel, P.
1965 "On the cost of being a Negro." Pp. 727-43 in E. O. Laumann, P. M. Siegel and R. W. Hodge (eds.), *The Logic of Social Hierarchies*. Chicago: Markham.
- Stolzenberg, R. M.
1975 "Education, occupation and wage differences between white and black men." *American Journal of Sociology* 81:299-323.
- Tannenbaum, A. S., B. Kavcic, M. Rosner, M. Vianello, and G. Wiesert
1974 *Hierarchy in Organizations*. San Francisco: Jossey-Bass.
- Thompson, V. A.
1961 *Modern Organizations*. New York: Knopf.
- Thurow, L. C.
1969 *Poverty and Discrimination*. Washington, D.C.: Brookings Institution.
- Winsborough, H. H., and P. Dickinson
1971 "Components of Negro-white income differences." *Proceedings of the Social Statistics Section of the American Statistical Association*, Ft. Collins, Colorado.
- Wright, E. O., and L. Perrone
1977 "Marxist class categories and income inequality." *American Sociological Review* 42:32-55.

CLASS POWER AND STATE POLICY: THE CASE OF LARGE BUSINESS CORPORATIONS, LABOR UNIONS AND GOVERNMENTAL REDISTRIBUTION IN THE AMERICAN STATES*

ALEXANDER HICKS, ROGER FRIEDLAND AND EDWIN JOHNSON

University of Wisconsin, Madison

American Sociological Review 1978, Vol. 43 (June):302-315

This paper investigates the impact of business and labor organizations upon governmental redistribution to the poor, or the extent to which government expenditures and revenues redistribute income to poor households. A cross-sectional analysis of 48 American states circa 1960 supports the propositions that large business corporations negatively affect governmental redistribution and that labor unions positively affect governmental redistribution. The analysis also supports past findings relating socioeconomic development, poverty, and the interaction of Democratic party strengths and cohesion and interparty competition to governmental redistribution. Findings suggest that redistribution to the poor by American state governments is a class issue, partially determined by conflicting class forces.

Typically, it has been argued that poor people in the United States are unorganized and thus denied access to the political system (Schattschneider, 1960). Excluded from the relatively high-wage unionized sector of the economy, the working poor are not organized politically via their place in production (O'Connor, 1973). As a result, the poor are forced into disruptive tactics outside the limits of American pluralism (Piven and Cloward, 1971; Alford and Friedland, 1975). In this view, the politics of poverty represent a desperate form of interest group politics which force political elites into cooptive and conciliatory concessions.

The thrust of our paper is that redistribution to the poor through government expenditure and taxation policies is a class issue. We argue that national corporations and labor unions are organizational bases of class power; that the distribution of these class organizations across states will

affect the pattern of redistribution to the poor; and that these effects will be consistent with a class conflict model of redistribution to the poor. Specifically, we hypothesize that the presence of national corporations will have a negative and labor unions a positive effect on state redistribution to the poor.

TOWARDS A CLASS MODEL OF GOVERNMENTAL REDISTRIBUTION

The potential impact of corporate and labor union power on redistribution to the poor has been virtually ignored in the empirical literature on state policy (for an exception, see DeLeon, 1973). In this paper we argue that corporate and labor union organizations have class interests in state government redistribution to the poor through public taxation and expenditure. To the extent that corporate or labor union organizations are concentrated in a state, the level of government redistribution to the poor is hypothesized to be more consistent with these class interests.

Interests generally shared between unionized and nonunionized segments of the working class can be conceptualized as class interests. Similarly, interests shared between corporations regardless of their industrial or market position can be conceptualized as class interests. Positive labor union interests in redistribution to

* The research for this paper was facilitated by the National Institute of General Medical Sciences Training Program in Methodology and Statistics (5-TO1-GM01526-09). The authors are indebted to Michael Aiken, Robert Alford, Howard Erlanger, Robert Hauser, Aage Sørensen, Maurice Zeitlin, Morris Zelditch, and anonymous reviewers, as well as the participants of the Social Organization Seminar and the Methodology Seminar at the University of Wisconsin, Madison, for helpful criticisms, comments and suggestions.

the poor might be expected to be minimal given the concentration of labor unions in the most capital intensive, highly productive, monopolistic sectors of the economy, and the concentration of the working poor in labor intensive, less productive, more competitive sectors (O'Connor, 1973). However, while the material condition of the poor is not relevant to the immediate collective bargaining interests of the relatively high-income union membership, it is critical to the political power of organized labor which makes these immediate gains possible. First, organized labor's political power is contingent upon the strength of the Democratic party whose electoral success in turn requires the political mobilization and votes of the nonunionized working class and dependent surplus populations (Greenstone, 1969). Second, labor unions are interested in protecting the private incomes of low income families through progressive taxation and supplementing them through public expenditures to increase the floor under the union wage. Third, labor unions are interested in using the public purse to protect and supplement the incomes of their own membership. Because taxation and social service programs are often organized in universalistic terms, nonunionized members of the working class are also likely to benefit. This is more than a trickle-down effect. On the contrary, because many of the poor are in fact exunion members who are retired, disabled or laidoff, much labor union interest in the poor is closely tied up with union members.

Corporate interests in redistribution to the poor, on the other hand, are negative. First, corporate political power is neither tied to the Democratic party nor is it greatly contingent upon electoral mobilization (Alford and Friedland, 1975). Second, corporations continue to be interested in the availability of low-income labor, particularly where open-shop legislation and routine, low-skill production make this possible. Third, corporate elites are interested in protecting their profitability and income both by minimizing the progressiveness of the tax structure and by securing the maximum share of public expenditures for programs which

are indirectly productive. High levels of redistribution to the poor involve both relatively progressive taxation and high levels of government expenditures which are not even indirectly productive.

Patterns of corporate and labor union elite behavior suggest that these interests are salient political concerns. National labor union elites are as ideologically supportive of welfare expansion and income equalization programs as national corporate elites are opposed (Barton, 1975). Labor unions are perhaps more effective in securing broad redistributive legislation benefiting the entire working class than they are in securing regulatory changes that benefit the more particularistic concerns of the union (Lowi, 1969). State-level case studies indicate strong, active corporate opposition to redistributive policy changes (see Fischer, 1974; Lockard, 1959). At the city level, labor unions have played an important role in organizing the Democratic vote of the working and dependent poor (Greenstone, 1969).

The resources that labor union and corporate organizations have to translate these class interests into policy are considerable. The potency of corporate elite political participation in state governments is intensified by their well-financed lobbyists (Schattschneider, 1960; Ziegler, 1969), monopolies over information and expertise required by government (McConnell, 1966), control over high-income career paths out of the public sector, and by their discretionary control over investment decisions upon which the efficacy of government policy is dependent (Haveman and Hamrin, 1973; Friedland, 1977a). Beyond its own lobbying apparatus, the political participation of labor union elites is buttressed by their control over an extensive apparatus for political mobilization of voters at all levels of government (Greenstone, 1969). Thus the very presence of corporations and labor unions which command concentrated political and economic resources may structure the range of political coalitions, issues and public policies without high levels of participation by their elites.

We have specified the preredistributive interests of labor unions and the antiredistributive interests of corporations, and the

considerable political resources these organizations can wield in the determination of public policy. Consequently, we hypothesize that *a strong labor union presence in a state will have a positive effect upon governmental redistribution to the poor, while a strong corporate presence in a state will have a negative effect upon governmental redistribution to the poor.*

Past Research Findings

Theory and research concerning the causes and consequences of government income redistribution through public expenditure and taxation have developed rapidly over the past two decades (e.g., Key, 1956; Dawson and Robinson, 1963; Lowi, 1964; Cutright, 1964; 1967; Fry and Winters, 1970; Hofferbert, 1972; Jackman, 1974; 1975; Wilensky, 1975). Such policies have been argued to be important for societal integration, business-cycle stability and long-term economic growth (see Shonfield, 1966:3-18; Galbraith, 1958:115-20; Samuelson, 1973:805-6). Recent economic research on the United States, Great Britain and Sweden indicates that the net redistributive impacts of government spending and taxing upon poverty and income inequality are quite substantial (see Reynolds and Smolensky, 1974; 1977; Nicolson, 1974; Franzen, 1975).¹

Quantitative research on the determinants of governmental redistribution consistently has found that socioeconomic development (e.g., per capita G.N.P. and median school years), redistributive need (e.g., percent poor), and governmental structure (e.g., executive centralization) significantly influence redistributive policies (see Fry and Winters, 1970; Fried, 1975; Jackman, 1975; Peters, 1974; Wilensky, 1975). The strength of rela-

tively left parties and the degree of unionization sometimes have been found to influence redistributive policy in Western European nations (e.g., Peters, 1974; Fried, 1973).² For the case of the United States, the Democratic party's electoral and office-holding strength has been found to have a positive impact upon redistributive policy outcomes under conditions of high party competition and cohesiveness (Hicks et al., 1976).

DEVELOPMENT, POVERTY, PARTY AND GOVERNMENTAL REDISTRIBUTION

Such aspects of socioeconomic development as income, education, industrialization and urbanization consistently have been found to affect redistributive policy outputs (e.g., Dawson and Robinson, 1963; Cutright, 1964; Lineberry and Fowler, 1967; Jackman, 1974; Peters, 1974; Wilensky, 1975). The enlargement of discretionary resources, the decline of kinship-based forms of social security, and the liberalizing effects of increased education are a few of the reasons that have been given for the tendency for government redistribution to increase with greater socioeconomic development.

The level of poverty, as measured by such variables as percent poor, also has been found to have consistently positive effects on redistribution (e.g., Fry and Winter, 1970; Fry, 1974; Wilensky, 1975). To what extent such effects are to be attributed to the social control and office-holding interests of governmental elites, to humanitarian responses of governmental personnel, or to perceived human needs has not been established. Whatever the reason for poverty's redistributive effects, past research does suggest that public expenditure will be more redistributive in jurisdictions with a large proportion of poor people.

Political scientists and sociologists have focused considerable attention upon the relation of party and electoral variables to public policies (see Dawson and Robin-

¹Reynolds and Smolensky (1974; 1977) find that the net impact of the fiscal activities of all levels of government in the United States has tended, over the past three decades, to augment the final incomes of the prefiscal poor by about 100%, and to decrease overall income inequality by more than 25%. Muzgrave et al. (1975) have generated congruent findings for the United States, as have Nicolson (1974) for Great Britain and Franzen (1975) for Sweden.

² Findings on the relation of Left party strength to redistributive policies in Western Europe are not all consistent; they vary from unit to unit and period to period (see Fried, 1975).

son, 1963; Cutright, 1964; Hofferbert, 1972; Jackman, 1974). Until recently, most evidence on the impact of electoral variables, such as the strength of the Democratic party and interparty competition, upon governmental redistribution suggested they were inconsequential (see Dawson and Robinson, 1963; Fry and Winter, 1970; Hofferbert, 1972). However, Hicks et al. (1976) found that Democratic party strength, Democratic party cohesion, and interparty competition interact with one another to generate more redistributive state policy. More precisely, they concluded that Democratic party strength has a positive effect on the level of state redistribution to the poor to the extent to which Democratic party cohesion and interparty competition are also simultaneously high in a state. The rationale is that Democratic party strength denoted by office-holding is not translated into progressive, redistributive legislation without the complementary *capacity* of party cohesion and the *imperative* of interparty competition. This hypothesis will be tested here in the context of new controls for corporate and union presence.

In this section, we have focused upon those variables in the state policy literature likely to affect governmental redistribution. However, a number of additional variables in the literature have been argued to influence governmental redistribution. The majority of such variables are referred to in Fry and Winters (1970) and Booms and Halldorson (1973). A number of variables, most of them drawn from these two sources, will be employed for control purposes in our statistical analysis (see below).

METHODOLOGY

Theory and Technique

Our use of an additive regression model to assess the impact of labor union and corporate presence on redistribution to the poor makes a series of implicit theoretical assumptions. First, holding corporate presence constant, we estimate the effect of labor union presence on redistribution. The implicit conception of power is additive, not interactive. We do not measure

directly class conflict, or even the relative organizational presence of corporations and labor unions. We treat a class power as a relation of the class to the state but not as a relationship between classes. Second, by assuming conflicting corporate and labor union interests in redistribution to the poor, and by arguing that cross-sectional variation in redistributive policy is partially and additively determined by class organization, we approach a pluralist theory cum class actors. Redistribution to the poor is determined by a vector sum of pressures by organized interests, each of which can have some independent access to the state. Third, we do not specify any functional role for the state, except as it transmits the power of organized interests into policy. The model specifies no variable capacity for states to adjust to general systemic problems generated by social disruption or economic contradictions, which may vary across states. Finally, the structure of the state is taken to be theoretically nonproblematic in that it is assumed not to affect corporate and/or labor union political power vis-à-vis redistributive policies to the poor across different U.S. states.

Units of Analysis

The American states have been a principal domain for the testing of hypotheses on determinants of redistributive policies (e.g., Dawson and Robinson, 1963; Fry and Winters, 1970; Booms and Halldorson, 1973; Hofferbert, 1972). This appears to be the case for a number of reasons. Substantively, the states are important U.S. political jurisdictions. They have authority over the bulk of public assistance and unemployment and workman's compensation programs; and they have, on the average, raised nearly a quarter of all governmental revenues and spent nearly a third of total government expenditures in the United States during the past three decades (see Sharkansky, 1972). Methodologically, states are rich in data from census sources and past study surveys; they are unusually free from problems of comparability which plague much comparative research; they are numerous enough for the application of fairly sophis-

ticated quantitative techniques; and they are few enough for economical quantitative research.

Measurement and Data

For our measure of state redistribution to the poor, we used a measure of the extent to which the total spending and taxing activities of state governments progressively redistribute income to poor households. This measure of state redistribution for 1961 was constructed by Booms and Halldorson (1973) with data from the U.S. Department of Labor, Bureau of Labor Statistics (1965), Bishop (1967), and the U.S. Department of Commerce, Bureau of the Census (1963). It resembles a number of other indices constructed during the past decade to measure governmental redistribution (e.g., Bishop, 1967; Reynolds and Smolensky, 1974). It operationally defines governmental redistribution as the ratio of estimated net expenditure benefits to estimated net revenue burdens for households with incomes of less than \$4,000. Values on this index range from a minimum value of 1.54 for South Dakota to a maximum value of 6.00 for Massachusetts. (See Table 1 and Booms and Halldorson [1973] for further information on this measure.)

South Dakota's 1.54 value on the redistribution index means that the average poor family received \$1.54 in public goods and money transfers for every dollar paid in taxes and user charges. A few figures

on transfer payments and tax schemes for this scale's extreme cases may serve to increase the reader's feel for the index. In South Dakota in 1961, \$46 was paid out in public assistance payments for every person below the federal poverty line for that year; the corresponding figure for Massachusetts was \$171. In the same year, the maximum annual unemployment payment for an individual was \$816 in South Dakota and \$1,350 in Massachusetts. Similarly, the mixes of (regressive) sales and (progressive) personal income tax revenues were strikingly different for these two states. In South Dakota in 1961, 61% state-raised revenues were raised by means of sales taxes, and 0% by means of personal income taxes. Corresponding figures for Massachusetts were 34% and 30%. (See U.S. Department of Commerce, 1963.)

Corporate and labor union presence were operationalized by the existence of their administrative and/or policy-making centers in a state. The presence of national corporate and labor union centers in a state indicates a variety of things. First, it indicates the presence of capitalists and working-class elites within the state who are available for state-level political participation. Second, it indicates the presence of organizational actors and organizational units of political representation in the state. Third, it indicates the size and diversity of the organizational base for class political action should it occur.

Table 1. Booms and Halldorson's Redistribution Index Values for 48 States,* 1961

6.004	Massachusetts	2.995	Pennsylvania	2.184	Wyoming
5.904	Illinois	2.958	Oregon	2.166	Delaware
5.804	California	2.785	Minnesota	2.128	Nebraska
5.385	Colorado	2.774	Nevada	2.126	North Carolina
4.700	Connecticut	2.653	Oklahoma	2.126	Florida
4.680	Washington	2.590	West Virginia	2.123	Louisiana
4.348	New York	2.559	Utah	2.036	Arkansas
4.342	New Jersey	2.522	Wisconsin	1.971	New Mexico
3.900	Ohio	2.522	Georgia	1.910	Montana
3.694	Missouri	2.482	Iowa	1.908	Vermont
3.679	Idaho	2.435	Arizona	1.845	New Hampshire
3.549	Michigan	2.398	Kentucky	1.740	South Carolina
3.501	Kansas	2.304	Alabama	1.728	North Dakota
3.375	Rhode Island	2.216	Tennessee	1.661	Virginia
3.372	Maryland	2.211	Maine	1.574	Texas
3.348	Indiana	2.198	Mississippi	1.537	South Dakota

* Mean: 3.172; standard deviation: 1.226.

For indicators of corporate presence, measures of the presence of policy-making and administrative centers of large financial and nonfinancial corporations are used. Items employed include the number of headquarters of the top 500 manufacturing corporations and the top 50 commercial banks located in a state, and the number of Domhoff's (1970) upper class clubs located in a state (see *Fortune*, 1961a; 1961b; Domhoff, 1970). Headquarters were emphasized and corporate plants ignored, because headquarters are loci for political activity, while plants are not, due to limited political autonomy (see Friedland, 1977a; Aiken, 1970:396-7).

A principal components analysis with varimax rotation was run on the corporate items to test for their unidimensionality and convergent validity.³ Only one principal component with an eigenvalue greater than one emerged from the analysis; and all thirteen corporate items loaded highly on this principal component.⁴ The factor score for the component was used to scale our measure of corporate presence.

³ The discriminate validity of this measure and of our measures of union presence, Democratic cohesion and interparty cohesion were assumed on theoretical grounds. For example, some items for the corporate presence measure correlated more highly with certain union presence items than with certain other corporate presence items. Still, they were regarded theoretically as indicators of corporate presence, which was expected to have impacts on governmental redistribution opposite to those of union presence. This expectation, as we shall see, was confirmed by our analysis.

⁴ The principal component for corporate presence explained 87.3% of the common variance in its constituent items. The thirteen items included numbers of headquarters in a state of the largest 50, 200, 500 and 1,000 manufacturing corporations and of the largest 50 banks and the largest 50 insurance companies (*Fortune*, 1961a; 1961b; 1963); the number, number of personnel of, and value added by administrative and auxiliary enterprises (Department of Commerce, Bureau of the Census, 1966); the numbers and assets of corporate foundations headquarters in a state (Department of Commerce, Bureau of the Census, 1963); the assets of all financial corporations headquartered in a state (U.S. Department of Commerce, 1963); and the number of Domhoff's (1970) upper class clubs located in a state. Loadings ranged between .985 and .843 and had a median value of .934. Only one principal component with an eigenvalue of 1.0 emerged from the analysis of these items. All components referred to in this paper have been constructed using varimax rotation.

For indicators of the presence of labor unions we used measures of the number of both union headquarters and locals. Union locals were employed because of their political autonomy (see Greenstone, 1969). Union membership seemed an appropriate indicator of union political resources because of the extensive union apparatus for member political mobilization.⁵ Principal components analysis was used to establish convergent validity and construct a scale for union presence.

To measure socioeconomic development, we used data on the most common development variables and measures in the American state policy literature. These were median personal income, median years of education, and percent of state population residing in cities with populations greater than 2,000. Several measures were then added to these to broaden the range of development measures. The percentages of state populations residing in central cities and SMSAs were added to complement the measure of urbanization noted above. Per capita value added in manufacturing was used as an additional measure of industrialization. Percent white-collar was introduced as a crude measure of socioeconomic status. In order to achieve a parsimonious mea-

⁵ One principal component emerged from the analyses of the four union presence items: (1) number of AFL-CIO members (in a state) (.955); (2) number of union headquarters (.925); (3) number of union locals (.923); (4) number of AFL-CIO affiliated locals listed in a state (.853). This component explained 83.7% of the common variance in the items. Loadings are listed above after items in parentheses. Data for these items are taken from U.S. Department of Commerce (1963) and U.S. Office of Labor-Management and Welfare-Pension Reports (1960). Data for the corporate and union presence measures are not standardized on the size of state (e.g., population, economically active population, etc.) because such standardization would control away corporate and union policy effects entailing the mediation of corporate and union leverage over the national political system (e.g., national parties, congressional committees, federal agencies). Thus, nationally mediated effects of a state's large business corporations on unions would be expected to be larger for, say, a five million population state with 500,000 union members, headquarters, than for a one million population state with 100,000 union members. In order to control for size effects other than corporate or labor presence effects, population size will be utilized below as a controlled variable.

surement of development, all items were entered into a principal components analysis. From this, two clear components emerged: socioeconomic development and urbanization-industrialization. Both will be used as development variables.⁶

Poverty is measured by the percentage of state households with incomes of less than \$3,000 in 1959. This measure of poverty has been the most commonly used measure of poverty in the state literature on redistributive policies. (See U.S. Department of Commerce, 1963, for source.)

The three party variables—Democratic strength, Democratic cohesion, and interparty competition—were also measured using separate principal components analyses to establish their convergent validity and to generate scales. The principal component for Democratic strength includes items on the average strength of the Democrats in terms of percentages of gubernatorial votes and legislative seats (1948 through 1962) and an item on the number of successive two-year periods during the same period marked by simultaneous Democratic possession of the governorship and Democratic majorities in all legislative houses. The items for interparty competition include two measures of absolute deviations from fifty-fifty party splits in voter support and legislative control (both averaged over the period 1948–1962). They also include two measures of interparty turnovers in elective office holding (1948–1962). The component for Democratic cohesion contains three items, two measures of Democratic voting cohesion and a measure of partisan conflict in state legislatures.⁷ (For a

theoretical and operational elaboration of these measures, see Hicks et al., 1976).

Eleven additional variables were incorporated into our statistical analysis as controls in order to check and correct for possible specification bias due to omitted variables. Three sorts of variables were selected. These are presented here in terms of the criteria used for their selection. First, those variables that had partial correlations in Fry and Winters (1970) and standardized regression coefficients in Booms and Halldorson (1973) which were greater than 0.20 in absolute value were used. These are income inequality, political participation, governor tenure potential, legislative professionalism and policy innovation, and percent of the labor force employed in nonagriculture. Second, Democratic strength and cohesion and interparty competition—the three variables composing the party variable interaction—were used. Third, population size (a possible source of spuriousness in the estimates for corporate and union presence) and a dummy variable for the South (a proxy for long-term, institutionalized one-party politics, etc.) were used. The procedures which we used to incorporate these variables into our analysis as controls are discussed below.⁸

Model and Techniques

Our model may be expressed in the form of the following equation:⁹

ance in its three items. The principal component for interparty competition explains 64.5% of the common variance in its four items. For further information on these measures, see Hicks et al. (1976).

⁸ Fry and Winters (1970) and Booms and Halldorson (1973) were our main sources of control variables because these studies alone approached exhaustive use of the principal operationalized variables in the state policy literature and because each of these studies used measures of governmental redistribution. See Fry and Winters (1970) for theoretical and operational discussions of variables in our first set of controls. Measurement of three party controls is discussed above and in Hicks et al. (1976). Population was measured as the natural logarithm of 1959 population figures from the U.S. Department of Commerce (1963). South was defined to equal the ten state region composed of Alabama, Arkansas, Georgia, Florida, Louisiana, Mississippi, North Carolina, South Carolina, Texas, and Virginia.

⁹ Our model might be expressed in the form of the following two equations:

⁶ The socioeconomic development principal component explained 81.5% of the common variance in measures of median household income (.927), median years of schooling (.884), and percent of the labor force in white-collar occupations (.881). (The three-digit numbers in parentheses are component loadings.) The urbanization-industrialization component explained .95% of the common variance in the percentage of state population in noncentral city portions of SMSAs (.889), the percentage of non-SMSA state population in cities with greater than 2,000 inhabitants (.976), and per capita value added in manufacturing (.989).

⁷ The principal component for Democratic strength explains 90.2% of the common variance in its three items. The principal component for Democratic cohesion explains 65.5% of the common vari-

$$\begin{aligned} \text{Redistribution} = & a - b_1 \text{ Corporate presence} \\ & + b_2 \text{ Union presence} \\ & + b_3 \text{ Poverty} \\ & + b_4 \text{ Socioeconomic development} \\ & + b'_4 \text{ Urbanization-Industrialization} \\ & + b_5 (\text{Democratic strength} \times \text{Democratic cohesion} \times \text{Interparty competition}) \\ & + b_6 X_6 + \dots + b_n X_n + e, \end{aligned}$$

where X_6, \dots, X_n are control variables.

$$\begin{aligned} \text{(i) Redistribution} = & a - b_1 \text{ Corporate presence} + b_2 \text{ Union presence} \\ & + b_3 \text{ Poverty} + b_4 \text{ Socioeconomic development} \\ & + b'_4 \text{ Urbanization-Industrialization} \\ & + b^*_5 \text{ Democratic strength} \\ & + b_6 X_6 + \dots + b_n X_n + e, \end{aligned}$$

where

$$\text{(ii) } b^*_5 = b'_5 + b_5 \text{ (Democratic cohesion} \times \text{interparty competition), } b_1 \text{ is less than zero, } b_2, b_3, b_4, b'_4, b^*_5 \text{ and } b_6 \text{ are greater than zero, and } X_6, \dots, X_n \text{ are control variables.}$$

If the values of b^*_5 are greater than zero, then Democratic strength positively affects governmental redistribution. If b_5 is greater than zero, then the magnitude of these effects varies directly with the product of interparty competition and Democratic cohesion.

Substituting $b'_5 + b_5$ for b^*_5 in equation (i) yields the equation in the text, *plus* the terms $b'_5 \times$ Democratic party strength.

This equation contains all of the parameters of equations (i) and (ii), excepting b^*_5 and b'_5 . Values for b^*_5 may be computed readily from equation (ii), given estimates for b'_5 and b_5 . b'_5 can be estimated by inclusion of Democratic party strength in this equation. Consequently, we can estimate the model of equations (i) and (ii) and test over hypotheses by estimating the parameters of the equation in the text and running significance tests on them.

It should be noted that *equation (ii) may be read as interpreting the interaction of Democratic strength, Democratic cohesion and interparty competition* as a series of effects of Democratic strength which are contingent for their magnitudes upon levels of (Democratic cohesion \times interparty competition). Also, the *exposition* of results for b^*_5 will be made simpler and clearer if the variables involved are expressed in scales which have real zero points. For example, an effect of Democratic strength for a scaled value of the product of Democratic cohesion and interparty competition equal to zero will be more effectively communicated if the scale zero value for this product connotes a situation of *no* cohesion and *no* party competition (i.e., a real zero) rather than, say, a mean value of the product variable which happens to be scaled equal to zero. The scales of Democratic strength, Democratic cohesion and interparty competition, which are in standard deviation units, all originally had scale means equal to

Analysis was performed in three rounds. First, the equation was estimated without the eleven control variables. Second, the equation was reestimated after ejection of those variables which did not have significant effects upon governmental redistribution at the .05 level, as indicated by results of the first round of estimation.¹⁰ Third, tests were run for significant effects of the eleven controls by means of a stepwise regression procedure for which a 0.05 level of significance was designated as a criterion for the incorporation of controls into our model.¹¹ Ordinary least squares was used for estimation.

zero and negative minimum values. Since the minimum value cases for each of these variables quite clearly constituted cases with approximate real zero values on the variables, it was easy to transform scale zeros into approximate real zeros by simply subtracting the minimum value of each of the original scales from each of the original scales (i.e., zero). This procedure anchors the zero value of our final scale for interparty competition to the case with minimum interparty competition (Louisiana); it anchors the zero value of Democratic cohesion to Louisiana as well; and it anchors the zero value of Democratic strength to Vermont.

¹⁰ Our decision to use tests of statistical significance for the analysis of data on an apparent population is based on the arguments and assumptions presented by Camilleri (1970) and Hagood (1970).

¹¹ "Stepreg I," a stepwise regression program based on Efroymson (1960) was used. The explanatory variables of Figure 1 were entered into the programs as base variables; the control variables were entered as free variables. A forward stepwise selection procedure was used. The 0.10 probability level was designated for determining when a free variable would enter or be ejected from a model at each step. Since hypotheses are one-tailed, a .05 probability level is the one effectively employed. The hypotheses employed for the control variables were that income inequality and South are negatively related to governmental redistribution and that all of the other controls are positively related to governmental redistribution. Hypotheses for control variables are based upon theory and results presented in Fry and Winters (1970) and Booms and Halldorson (1973) for all controls but population size and South. Population size was hypothesized to have a positive impact upon governmental redistribution because of the assumed greater leverage of larger states over federal grants. South was hypothesized to have a negative impact upon governmental redistribution because of its conservative political tradition. Because the order in which free variables are entered into a stepwise regression procedure may alter final results, all stepwise analyses were repeated with an inversion of the original ordering of free variable. Results were identical using either ordering.

Table 2. Matrix of Zero-Order Correlations for Variables of 48-State Model of Government Redistribution

	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇
X ₁ Corporate presence	1.00						
X ₂ Union presence	.85	1.00					
X ₃ Poverty	-.30	-.40	1.00				
X ₄ Socioeconomic development	.34	.42	-.92	1.00			
X ₅ Dem. strength x Dem. cohesion x interparty competition	.04	.08	-.48	.44	1.00		
X ₆ Gov. tenure potential	.28	.29	-.41	.59	.25	1.00	
X ₇ Policy innovation	.58	.72	-.62	.45	.30	.23	1.00
Y Redistribution	.42	.62	-.56	.64	.48	.47	.70

T-statistics were used for all statistical tests. All tests were one-tailed. In the section on findings, results for the end of the third round of estimation will be emphasized.

Findings

Our hypotheses regarding the effect of labor union and corporate presence are supported by our regression analyses (see Table 3 and Figure 1).¹² Labor union presence has positive effect upon governmental redistribution, while corporate presence has a negative effect upon governmental redistribution. The effect of corporate presence upon governmental redistribution is $-.390$ in raw metric terms and $-.329$ in standardized terms (see \hat{b}_1 and $\hat{\beta}_1$ in Table 3). The effect of union presence is $.575$ in raw metric terms; it is

$.484$ in standardized terms (see \hat{b}_2 and $\hat{\beta}_2$ in Table 3).¹³

Two of the eleven control variables are found to influence significantly governmental redistribution. These are governor tenure potential, and policy innovations (see below). Our model's fit with the data is good. When our model includes governor tenure potential and policy innovations, it has a coefficient of determination corrected for degrees of freedom of $.71$. When it excludes these two controls, it has a coefficient of determination corrected for degrees of freedom of $.65$ (see Table 3). Our presentation of findings is confined to those for the model which includes the two controls since estimates do not vary substantially between the two variants of our model and since the longer model reveals the shorter one to be misspecified (see Table 3).

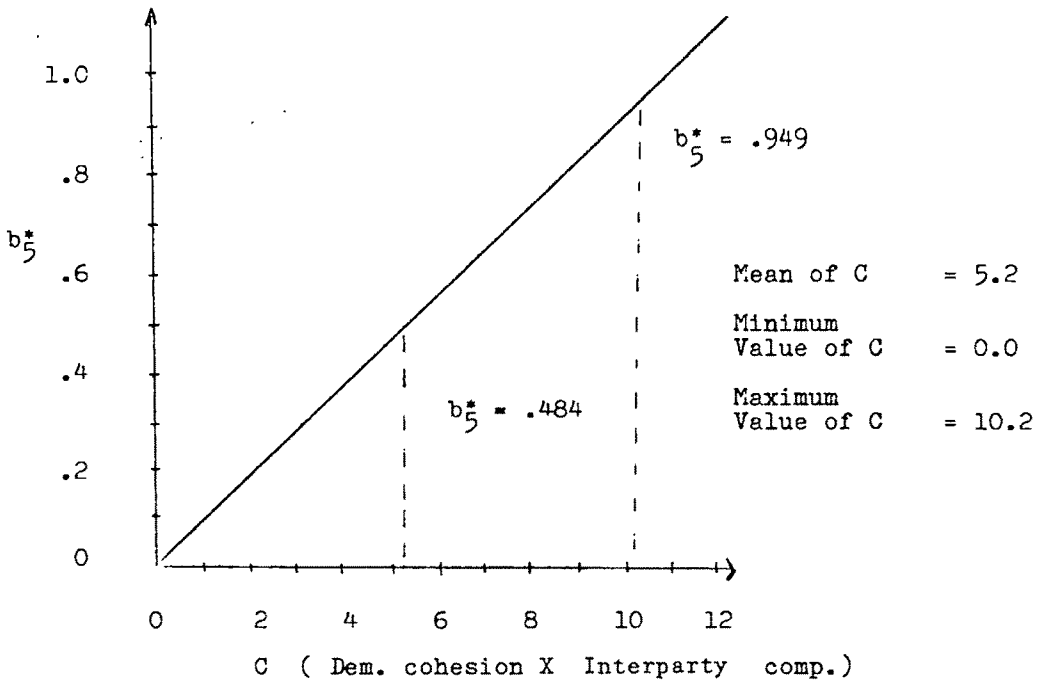
Poverty and socioeconomic development both have positive impacts upon governmental redistribution. The effects of poverty are $\hat{b}_3 = .068$ and $\hat{\beta}_3 = .582$. Those of socioeconomic development are $\hat{b}_4 = .755$ and $\hat{\beta}_4 = .635$ (see Table 3).

The party variable interaction term has a positive impact upon governmental redistribution ($\hat{\beta} = .276$). Under the model of

¹² A look at Table 2 will quickly reveal that many of the regressors in our final model are highly correlated. More particularly, the zero-order correlations between poverty and socioeconomic development ($r = -.92$) and between corporate presence and union presence ($r = .85$) may suggest serious difficulties to some readers.

However, the presence of imperfect multicollinearity becomes problematic only to the extent that it renders standard errors of regression parameter estimates unacceptably high (Johnston, 1972:1963-4). This is not the case for this paper's analyses. In these, despite possibly inflated standard errors due to multicollinearity, statistical significance is achieved for all principal tests on regression parameters (see Table 3). As Klein (1962:50) notes for the case of two regressor regressions and as Theil (1971: 165-7, 163-4) demonstrates for the general multivariate case, the potentially destabilizing effects of multicollinearity in a parameter estimate \hat{b}_{yx_1} may be offset either by (a) a high R^2 for the equation, or (b) a high sum of squares for x_1 .

¹³ Also, we cannot reject the null hypothesis $\beta_1 + \beta_2 = 0$ at the $.05$ level. That is, the redistributive impact of union presence is not significantly greater than the antiredistributive one of corporate presence. The t-statistic for the null $\beta_1 + \beta_2 = 0$ equals $(\hat{\beta}_1 + \hat{\beta}_2) / \sqrt{\text{var}(\hat{\beta}_1) + \text{var}(\hat{\beta}_2) + 2 \text{Cov}(\hat{\beta}_1, \hat{\beta}_2)}$. The formula for the t-test for the null hypothesis that $\beta_1 - \beta_2$ is $\hat{\beta}_1 - \hat{\beta}_2 / \sqrt{\text{var}(\hat{\beta}_1) + \text{var}(\hat{\beta}_2) - 2 \text{Cov}(\hat{\beta}_1, \hat{\beta}_2)}$. Tests for sums and differences of coefficients are all done with standardized coefficients (and on unstandardized coefficients with variables standardized on their own variances).



Note: Redistribution = $-2.918 - .390$ Corp. Pres. + $.575$ Union Pres. + $.068$ Poverty + $.756$ Dev't + b_5^* Dem. Strength + $.222$ G.T.P. + $.0005$ Innovations + e , where $b_5^* = .093$ (Dem. Cohesion \times Interparty Comp.)

Figure 1. Graph of b_5^* for Values of (Dem. Cohesion \times Interparty Competition)

equations (i) and (ii) (see footnote 9), the redistributive effect of Democratic strength is 0.00 where the product of Democratic cohesion and interparty competition is 0; it is .484 evaluated at the mean of this product; and it is .949 evaluated at the product's maximum value (see Figure 1).¹⁴

The control variable, governor tenure potential, has a significant positive effect upon governmental redistribution ($\hat{b}_8 = .222$, $\hat{\beta}_8 = .227$) in our regression analysis of 48 states (see Table 3). Governor tenure potential has been found to positively affect governmental redistribution on the assumption that (1) a governor's tenure potential, as an indicator of his power

over careers in state politics, is an important gubernatorial resource; and (2) the broad constituency support needs for gubernatorial election requires redistributive appeals and action from governors (see Fry and Winter, 1970).

The control variable innovations, an index of the overall rapidity with which state governments have adopted a broad range of policy innovations in the period 1870-1960, has a positive effect upon governmental redistribution. Because of its long-time coverage, it was regarded as an exogenous variable. Its effects are $\hat{b}_7 = .0005$ and $\hat{\beta}_7 = .392$. Policy innovativeness was hypothesized to positively affect governmental redistribution on the assumption that it tapped the "equitable, efficient, issue-oriented, progressive tradition in American politics" (Fry and Winter, 1970:518).¹⁵

¹⁴ For the exposition of findings regarding the interaction of party variables, we will focus our attention on unstandardized or raw metric coefficients (see, for example, Blalock, 1964:Chap. 4). Regarding the interpretation of these findings, it should be noted that the party variables, like all variables in the equation of the Model and Techniques section, are in standard deviation units. Information on the raw metric for governmental redistribution is displayed in Table 1.

¹⁵ Analytic units, such as states, which are certain to have many interrelations with one another and with the common systems (e.g., national) of which they are a part, are likely to entail errors in equations which are not independent across observations (see Duncan et al., 1961:129-60). In other words, they

Table 3. Parameter Estimates for Regression Analysis of State Governmental Redistribution *

	X ₁ : Corp. Pres.	X ₂ : Union Pres.	X ₃ : Poverty	X ₄ : Dev't.	X ₅ : Interaction	X ₆ : G.T.P.	X ₇ : Innovs.
$\hat{\beta}_1$	-.329	.484	.582	.635	.276	.227	.392
$\hat{\beta}_1$	(-.326)	(.742)	(.440)	(.704)	(.331)		
\hat{b}_1	-.390	.575	.068	.755	.093	.222	.0005
\hat{b}_1	(-.387)	(.882)	(.052)	(.837)	(.112)		
SE \hat{b}_1	.176	.210	.025	.251	.031	.089	.0002
	(.194)	(.201)	(.027)	(.274)	(.034)		
p	.016	.004	.003	.005	.002	.008	.004
	(.026)	(.000)	(.034)	(.002)	(.001)		
a=	-2.918	R ² =.714					
	(.194)	(.647)					

* Results for equations run without the control variables governor tenure potential (G.T.P.) and innovations (Innovs.) are presented in parentheses.

$\hat{\beta}_i$ =estimated standardized regression coefficient for variable $i=1,\dots,7$.

\hat{b}_i =estimated unstandardized regression coefficient for variable $i=1,\dots,7$.

SE(\hat{b}_i)=standard error of \hat{b}_i ; a=regression intercept; p=probability value of \hat{b}_i .

\bar{R}^2 =coefficient of determination corrected for degrees of freedom.

In conclusion to our presentation of findings, let us note that urbanization-industrialization and nine of the eleven control variables were *not* significantly related to redistribution. This means that variables such as voter turnout, South and population size did not have to be included in our final models in order to obviate specification error.¹⁶

may well entail areally correlated errors, a spatial analogue of autocorrelated errors in time-series. We shall use Geary's (1954:115-45) R statistic to test for areal association among contiguous units, a particular form of areal correlation. If Geary's R is not significantly different from zero, we may have some confidence that the errors in our equations for governmental redistribution are uncorrelated across observations (see Duncan et al., 1961:131-6). Geary's R was computed for the residuals for the seven variable models including governor tenure potential and policy innovations. Results of this test do not allow us to reject the null hypothesis that the errors for this model are really uncorrelated, and they increase our confidence in our findings. For this test $R = .42$ and the probability of R being greater than or equal to .42 was .663, assuming the normal distribution of R. See Duncan et al. (1961:131-6) for the definitions and test procedures associated with Geary's R.

¹⁶ Additional analyses were run under the assumption that the three party variables and governor tenure potentials are causally subsequent to the other variables in our model and might mediate (and control away) effects of the other truly exogenous variables. In these analyses we estimated reduced-form coefficients for these exogenous variables. These analyses and their results were not integrated into

Discussion

Our findings of significant positive effects of labor union presence and negative effects of corporate presence on governmental redistribution to the poor at the state level suggest, first, that labor unions and corporations are organizational bases of class power and, second, that redistribution to the poor is a class issue. That these findings control for electoral characteristics suggests that corporations and labor unions exert an influence upon state policies which is unmediated by electoral politics. Our findings of net effects of the interaction of Democratic party strength and cohesion and interparty competition

the main body of this paper because the results of these analyses did not differ significantly from the results of the analyses presented here. The reduced forms for corporate presence, union presence, socioeconomic development, poverty, and innovations are very similar to the structural coefficients for these variables estimated. One notable result stems from this fact: the party variables and governor tenure potential do not mediate substantially any effects of the other variables in our analysis. No multiequation model with direct and indirect effects was presented here because we have not yet formulated causal explanations for such potential intervening variables as Democratic strength. Innovations was treated as an exogenous variable in our estimation of reduced forms because it measures innovations going back from 1960 to 1870. Reduced-form estimates for our models are available upon request.

on state redistribution to the poor suggests that electoral politics indeed does make a difference where a more progressive political party has *both* the legislative capacity and the electoral imperative to pursue redistributive policies. That these party characteristics affect redistribution independently of corporate and labor union strength indicates that the party system provides the electorate with a partially autonomous means of influencing redistributive policy (see Hicks et al., 1976). Class politics and party politics both matter in the determination of redistributive policy. So, in consonance with much previous research, do socioeconomic development, poverty, governor tenure potential, and policy innovativeness (see Findings above).

Research is accumulating which indicates that class power operates on subnational policy making via the location of policy-making and administrative centers of national corporations and labor unions. Friedland (1977b) has found analogous additive positive and negative effects of corporate and labor union presence on central city responses to the War on Poverty. Research is being conducted on the historical development of corporate and labor union power (Hicks, forthcoming), as well as its impact on state policies which are more directly related to profitability and wage interests of corporations and labor unions (Johnson, forthcoming). A dozen years after Lowi (1964:707) wrote that "issues that involve redistribution cut closer than any others along class lines and activate interests in what are roughly class terms," systematic evidence of class power as a determinant of redistributive public policy is being generated.

The challenge is to begin to develop models of policy determination with appropriate methodologies which specify the relational nature of class power and the impact of state structure and public policy on class power (see Friedland et al., 1977).

REFERENCES

- Aiken, Michael
1970 "The distribution of community power: structural bases and social consequences." Pp. 487-525 in Michael Aiken and Paul E. Mott (eds.), *The Structure of Community Power*. New York: Random House.
- Alford, Robert R. and Roger Friedland
1975 "Political participation and public policy." *Annual Review of Sociology* 1:429-79.
- Barton, Allen
1975 "Consensus and conflict among American leaders." *Public Opinion Quarterly* 38:507-30.
- Bishop, G. A.
1967 *Tax Burdens and Benefits of Government Expenditures by Income Class, 1961, 1965*. New York: Tax Foundation.
- Blalock, Hubert M.
1964 *Causal Inference in Non-Experimental Research*. Chap. 4. Chapel Hill: University of Illinois Press.
- Booms, Bernard and James R. Halldorson
1973 "The politics of redistribution: a reformulation." *American Political Science Review* 77:920-33.
- Camilleri, Santo F.
1970 "Theory probability and induction in social research." Chap. 16 in Denton E. Morrison and Ramon E. Henkel (eds.), *The Significant Test Controversy: A Reader*. Chicago: Aldine.
- Cutright, Phillips
1964 "Political structure, economic development and national social security programs." *American Journal of Sociology* 70:537-50.
1967 "Inequality: a cross-national analysis." *American Sociological Review* 32:180-90.
- Dawson, Richard and James Robinson
1963 "Interparty competition, economic variables, and welfare politics in the American states." *Journal of Politics* 25:265-89.
- DeLeon, Richard E.
1973 "Politics, economic surplus, and redistribution in the American states: a test of a theory." *American Journal of Political Science* 4:769-81.
- Domhoff, G. William
1970 *The Higher Circles*. New York: Random House.
1972 *Fat Cats and Democrats*. Englewood Cliffs: Prentice-Hall.
- Duncan, Otis Dudley, Ray P. Cuzzort and Beverly Duncan
1961 *Statistical Geography*. Glencoe: Free Press.
- Efroymsen, M. A.
1960 "Multiple regression analyses." Pp. 191-203 in R. Ralston and H. Will (eds.), *Mathematical Methods in Digital Computers*. New York: Wiley.
- Fortune
1961a July:168-83.
1961b August:132-44.
1963 *Plant and Product Directory of the 1000 Largest U.S. Industrial Corporations*. New York: Time and Life.
- Fischer, Glen
1974 *The Politics of the Purse*. Urbana: University of Illinois Press.

- Franzen, Thomas, Lovgren Kerstein and Irene Rosenberg
 1975 "Redistributional effects of taxes and public expenditures." *Swedish Journal of Economics* 77:31-5.
- Fried, Robert C.
 1975 "Comparing urban policy and performance." Chap. 6 in Fred I. Greenstein and Nelson W. Polsby (eds.), *Handbook of Political Sciences*, Vol. 6. Reading: Addison-Wesley.
- Friedland, Roger
 1977a *Class Power and the City*. Ph.D. dissertation, Department of Sociology, University of Wisconsin, Madison.
 1977b "Class power and social control: the case of the war on poverty." *Politics and Society* 7:459-89.
- Friedland, Roger, Francis Fox Piven and Robert R. Alford
 1977 "Political conflict, urban structure and the fiscal crisis." Pp. 197-224 in Douglas Ashford (ed.), *Comparative Public Policy: New Approaches and Methods*; Sage Yearbook in Politics and Public Policy, 1977. New York: Sage.
- Fry, Brian R.
 1974 "An examination of the relationship between selected electoral characteristics and state redistributive efforts." *American Journal of Political Science* 18:421-31.
- Fry, Brian R. and Richard F. Winters
 1970 "The politics of redistribution." *American Political Science Review* 49:508-22.
 1974 "Examination of the relationship between selected electoral characteristics and state redistributive effort." *American Journal of Political Science* 18:421-31.
- Galbraith, John Kenneth
 1958 *The Affluent Society*. Boston: Houghton Mifflin.
- Geary, R.C.
 1954 "The contiguity ratio and statistical mapping." *Incorporated Statistician* 5:115-45.
- Greenstone, J. David
 1969 *Labor in American Politics*. New York: Knopf.
- Hagood, Margaret Jorman
 1970 "The notion of a hypothetical universe." Chap. 4 in Denton E. Morrison and Ramon E. Henkel (eds.), *The Significant Test Controversy: A Reader*. Chicago: Aldine.
- Haveman, Robert and R. D. Hamrin
 1973 *The Political Economy of Federal Policy*. New York: Harper and Row.
- Hicks, Alexander
 Forthcoming *The Political Economy of Redistribution: The Case of the American States between the Great Depression and the Great Society*. Ph.D. dissertation, Department of Sociology, University of Wisconsin, Madison.
- Hicks, Alexander, Roger Friedland and Edwin Johnson
 1976 "Democratic competitive mobilization." Unpublished paper, Department of Sociology, University of Wisconsin, Madison.
- Hofferbert, Richard I.
 1972 "State and community policy studies: a review of comparative input-output analysis." Pp. 3-72 in James A. Robinson (ed.), *Political Science Annual: An International Review*, Vol. 3. New York: Bobbs-Merrill.
- Jackman, Robert
 1974 "Political democracy and social equality: a comparative analysis." *American Sociological Review* 39:29-45.
 1975 *Politics and Social Equality*. New York: Wiley.
- Johnson, Edwin
 Forthcoming *The Political Economy of Public Policy: A Model of Organized Labor and State Economic Policy*. Ph.D. dissertation, Department of Sociology, University of Wisconsin, Madison.
- Johnston, John
 1972 *Econometric Methods*. 2nd ed. New York: McGraw-Hill.
- Key, V. O.
 1956 *Politics in the American States*. New York: Knopf.
- Klein, Lawrence Robert
 1962 *An Introduction to Econometrics*. Englewood Cliffs: Prentice-Hall.
- Kruskal, William
 1974 "The ubiquity of statistics." *American Statistician* 28:3-6.
- Lineberry, Robert L. and Edmund P. Fowler
 1967 "Reformism and public policy in American cities." *American Political Science Review* 61:701-16.
- Lockard, Duane
 1959 *New England State Politics*. Princeton: Princeton University Press.
- Lowi, Theodore
 1964 "American business and public policy." *World Politics* 16:677-719.
 1969 *The End of Liberalism: Ideology, Policy and the Crisis of Political Authority*. New York: Norton.
- McConnell, Grant
 1966 *Private Power and American Democracy*. New York: Knopf.
- Martin, Andrew
 1973 "The politics of economic policy in the United States." Beverly Hills: Sage.
- Musgrave, Richard, K. Case and H. Leonard
 1975 "The distribution of fiscal burdens and benefits." *Public Finance Quarterly* 2:259-312.
- Nicolson, J. L.
 1974 "The distribution and redistribution of income in the United Kingdom." Chap. 3 in Dorothy Wederburn (ed.), *Poverty, Inequality, and Class Structures*. Cambridge, Ma.: Cambridge University Press.
- O'Connor, James
 1973 *The Fiscal Crisis of the State*. New York: St. Martin's.
- Peters, Guy
 1974 "Income redistribution: a longitudinal analysis of France, Sweden and the United Kingdom." *Political Studies* 25:311-23.

- Pivan, Francis Fox and Richard A. Cloward
1971 *Regulating the Poor: The Functions of Social Welfare*. New York: Pantheon.
- Rao, Potluri and R. L. Miller
1971 *Applied Econometrics*. Belmont: Wadsworth.
- Reynolds, Morgan and Eugene Smolensky
1974 "The post-fisc distribution: 1961 and 1970 compared." *National Tax Journal* 27:515-30.
1977 *Public Expenditures, Taxes and the Distribution of Income; the U.S.: 1950, 1961, 1970*. New York: Academic Press.
- Samuelson, Paul Anthony
1973 *Economics*. 9th ed. New York: McGraw-Hill.
- Schattschneider, Elmer Eric
1960 *The Semi-Sovereign People, A Realistic View of Democracy in America*. New York: Holt.
- Schonfield, Andrew
1966 *Modern Capitalism*. London: Oxford University Press.
- Sharkansky, Ira
1972 *The Maligned States*. New York: McGraw-Hill.
- Theil, Henri
1971 *Principles of Econometrics*. New York: Wiley.
- U. S. Bureau of the Census
1966 *Census of Manufacturing*. Washington, D.C.: U.S. Government Printing Office.
- U.S. Department of Commerce
1950- *Statistical Abstract of the United States*.
1964 Washington, D.C.: U.S. Government Printing Office.
- U.S. Department of Labor, Bureau of Labor Statistics
1965 *Survey of Consumer Expenditures: 1960-1961*. Washington, D.C.: U.S. Government Printing Office.
- U.S. Office of Labor-Management and Welfare-Pension Reports
1960 *Register of Reporting Labor Unions*. Washington, D.C.: U.S. Government Printing Office.
- Wilensky, Harold
1975 *The Welfare State and Equality*. New York: Wiley.
- Zeigler, Luther Harmon
1969 *Lobbying: Interaction and Influence in American State Legislatures*. Belmont: Wadsworth.

MANUSCRIPTS FOR THE ASA ROSE SOCIOLOGY SERIES

Manuscripts (100 to 300 typed pages) are solicited for publication in the *ASA Arnold and Caroline Rose Monograph Series*. The Series welcomes a variety of types of sociological work—qualitative or quantitative empirical studies, and theoretical or methodological treatises. An author should submit three copies of a manuscript for consideration to the Series Editor, Professor Robin M. Williams, Jr., Department of Sociology, Cornell University, Ithaca, New York 14853.

DETERMINANTS AND BEHAVIORAL CONSEQUENCES OF PSYCHOLOGICAL MODERNITY: EMPIRICAL EVIDENCE FROM COSTA RICA*

MICHAEL ARMER
Indiana University

LARRY ISAAC
Florida State University

American Sociological Review 1978, Vol. 43 (June):316-334

A major goal of research on national development and underdevelopment is to specify and empirically test propositions drawn from theories emphasizing individual, societal, or international determinants. In this paper we examine a central premise of individual modernity theory that psychological modernity mediates the effects of background factors and directly affects individual behaviors thought to contribute to societal modernization. Several analytical tests were made using fifteen behaviors identified in the modernization literature and an index of modern behavior. Based on data collected from a stratified quota sample of 210 Costa Rican adult males, we (1) estimate separate ordinary least-squares regression equations of the behaviors as a function of psychological modernity and background variables (age, rural-urban residence, education, occupation and income); (2) perform a similar analysis after forming indexes of background and behavioral indicators; and (3) estimate a full structural equation model of eight behaviors incorporating measurement error of psychological modernity and allowing the disturbances in the equations to be correlated. The results indicate that psychological modernity has, in most cases, a negligible effect on behavior when measurement is assumed to be perfect, and nonnegligible effects in the direction predicted for three of the eight behaviors when the measurement error in psychological modernity is taken into consideration. After evaluating design limitations and potential objections, we note that psychological modernity is important in determining only a limited number of behaviors and generally adds little to explaining behavioral variations beyond objective background characteristics. We conclude that psychological modernity appears more as an interpretative construct (and that only in several cases) than a pervasive source of modern behavior. Implications for national development and additional lines of research are discussed.

For several decades the sociology of national development has been based on (1) social structural perspectives which emphasize organizational and institutional determinants of development and (2) social psychological perspectives which emphasize personality determinants of development. The social structural theorists, whether arguing from conflict or functionalist positions, have sought to explain increasing differentiation, complexity, and industrialization of societies in terms of changes in the social structures and

processes within society, while the social psychological theorists have concentrated on changes in motivations and values of individuals as forces of social change. Horowitz (1970), Coleman (1971) and others have argued persuasively that development depends on an interaction of both social conditions and personality factors, and that a more useful analytic approach involves a general model of change in which, for a given society, certain of these elements may be seen as deficient.

Both structural and psychological perspectives have been overshadowed in recent years by the advance of world-systems theory which has emphasized international structures and processes in explaining development and underdevelopment of societies (e.g., Wallerstein, 1974; Chase-Dunn, 1975; Robinson, 1976). World-system perspectives stand in sharp contrast to earlier ontogenetic theories by calling attention to determinants external to the society and its members as well as

* We are indebted to Jairo Arboleda and Charles Kuhlman for assistance in data collection and preparation, and to Carolyn Mullins for editorial assistance. We also gratefully acknowledge critical reviews of earlier drafts by Allan Schnaiberg, Sheldon Stryker, C. Milton Coughenour, and an anonymous referee. Support and sponsorship for data collection and processing was provided by UNESCO, the Costa Rican Ministry of Public Education, the University of Costa Rica Division of Human Sciences, and the Indiana University Center for Innovation in Higher Education and Wrubel Computing Center.

to alternative directions societal development may take, including change toward underdevelopment and dependency (Hechter, 1975).

The distinction between these three levels of determinants—societal, individual, and international—parallels and generalizes Portes's (1976) distinction between social differentiation, value enactment, and liberation from dependency perspectives in his perceptive review of the sociology of national development. Portes criticizes all three but clearly favors world-system theory over the first two.

The shift in attention to world-system determinants is a welcome change from previous preoccupation with societal and individual determinants of socioeconomic development and nation building. On the other hand, stressing the importance of neglected factors and processes at any one level should not be allowed to develop into a deterministic position which recognizes only those factors. The advance of theory and evidence at the international level does not negate the importance of determinants and conditions at the societal and individual levels. These perspectives are not mutually exclusive. Horowitz's and Coleman's comprehensive approaches to competing perspectives on social change have much to recommend it. The eventual goal of research on national development should be a general model encompassing major resources and transformation processes which contribute to national development or underdevelopment at all three levels.

Such integration and synthesis is not likely until gains are made in specifying, empirically testing, and refining theoretical propositions within each of these levels rather than by claims of the superiority or inferiority of explanations at one level or another. Portes has summarized major strengths and weaknesses of each of these perspectives. None has a monopoly on logic or explanatory power, and empirical testing of propositions from each perspective is extremely limited.

The present study is focused at the individual level of determinants considered important for national development. In particular, we examine evidence bearing

on one of the most popular and pervasive of the value-enactment perspectives, the individual modernity thesis as advanced by Inkeles (1969), Inkeles and Smith (1974) and others. Our purpose is not to add to the already abundant theoretical debate over the individual modernity thesis (or other perspectives) in the sociology of national development, but rather to bring to bear empirical evidence from one Third World country on an important hypothesis at the individual level. In so doing, we hope to help advance understanding of individual processes in national development and to encourage further empirical testing and refinement of individual, societal and international perspectives.

THE MODERNITY THESIS

Central to individual modernity theory is the assumption that psychological modernity leads to modern behavior, which contributes to or is necessary for modernization of societies (see Lerner, 1965; Peshkin and Cohen, 1967; Kahl, 1968; Inkeles, 1969; 1971; 1975; Chodak, 1973; Portes, 1976; Portes and Cobas, 1976):

Attitude and value change defining individual modernity are accompanied by changes in behavior precisely of the sort which we believe give meaning to, and support, those changes in political and economic institutions which lead to the modernization of nations. (Inkeles and Smith, 1974:312)

To date, there have been substantial theoretical criticisms but little empirical evidence to support or refute this assumption. Both proponents and critics have been less concerned with testing the *consequences* of psychological modernity for personal behavior and societal development than with measuring and explaining its sources (recent examples include Feldman, 1975; Smith and Inkeles, 1975; Gough, 1976; 1977; Cohen and Till, 1977; Suzman, 1977). Yet the justification of this continuing research for a sociological understanding of national development depends largely on whether the attitudes and values defining modernity do, in fact, produce behavior of the sort indicated by modernity theorists as necessary for modernization of society. Of course, a deter-

mination of whether the behaviors, in turn, contribute significantly to national development must await cross-national studies involving longitudinal survey data of representative population samples. In the present study, empirical evidence showing that psychological modernity explains modern behavior would indicate the importance of pursuing such longitudinal cross-national research, while evidence to the contrary would endorse the skepticism of Wallerstein (1976), Portes (1976) and others about the importance of modern values and attitudes for national development.

The necessity of investigating the consequences of individual modernity has been recognized by proponents as well as critics of the modernity thesis. "To question the social significance of modernity research is to inquire about the centrality of the particular set of attitudes and values studied in real problems of national development" (Holsinger, 1974:44). In a related vein, Inkeles (1976:129) acknowledges: "The case for the social consequences of individual modernity rests mainly on evidence showing that becoming more modern leads individuals to undertake new transformative social roles within their societies."

Some researchers profess interest in modern attitudes as a socialization phenomena accompanying national development; modernity for its own sake rather than for its potential contribution to national development. The present investigation has relevance for these socialization studies insofar as modern values and attitudes are thought to lead individuals to new social roles and behaviors of the sort mentioned in the modernity literature.

The primary focus of this study is on the national development implications rather than the socialization implications of individual modernity. If psychological modernity (1) is an important determinant of modern behavior or (2) mediates much of the influence of educational experience and other background (demographic and socioeconomic) factors on such behaviors, or (3) both, then modern value orientations gain significance in the study of national development. Such evidence would justify further research on the con-

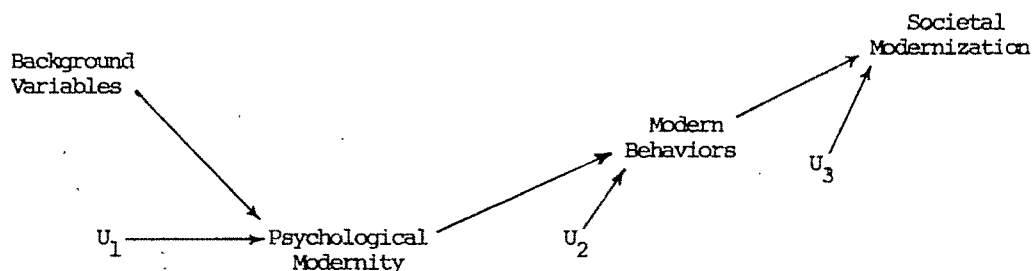
tribution of individual personality for the social, economic and political development of societies. In contrast, evidence that neither condition held would counter a major assumption of modernity theory, and would raise questions about the significance of modernity in the study of the role of individual factors in national development.

The individual modernity thesis of national development can be represented as a basic model of linear, additive effects shown in Figure 1, A. Background characteristics and experiences of individuals produce psychological modernity which is manifested in modern behavior which, in turn, contributes to societal modernization. Extraneous influences on the endogenous variables are represented by U's in the model.

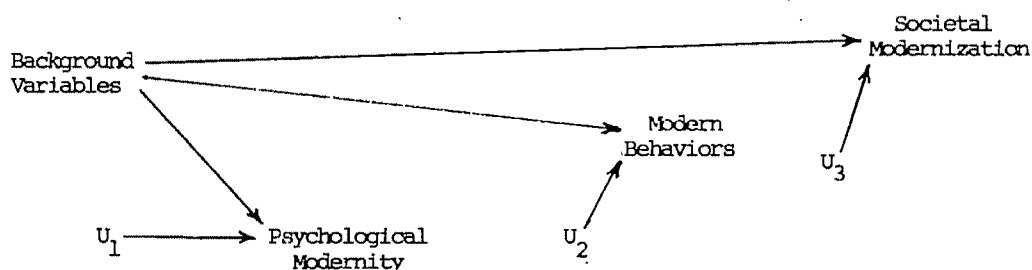
An alternative model of the relationship among these variables is proposed in Figure 1, B showing background factors exerting a direct effect on psychological modernity (as in the modernity thesis) but no important influence of psychological modernity on modern behavior or of modern behavior on social modernization. Rather, background factors have direct effects on all three endogenous variables and the significance of psychological modernity and behavioral modernity for societal modernization is nil.

The focus of most previous research has been in the initial link between background factors and psychological modernity. We propose to examine empirically not just this direct link, but also the possibility of direct and indirect links between background factors and modern behaviors (Figure 1, C). If the traditional model is accurate, then psychological modernity should exert its own substantial net effect on modern behaviors, mediate much of the influence of background factors, or both. If the alternative model is accurate, then the direct effect of psychological modernity and its mediation of background influences should be negligible.

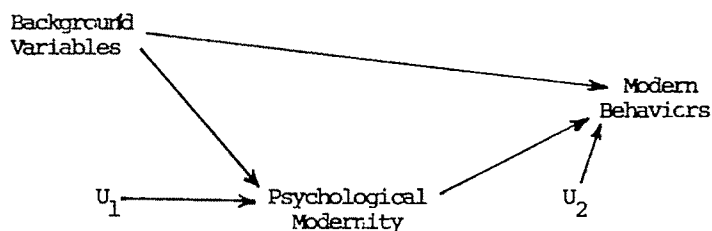
Lack of longitudinal cross-national data precludes an assessment of the link between psychological modernity and societal modernization, or between modern behaviors and societal modernization.



A. Traditional model of psychological modernity effects.



B. Alternative model of psychological modernity noneffects.



C. Model estimated in this research.

Figure 1. Models of Psychological Modernity Effects and Noneffects

Rather, we deal with what Inkeles and Smith (1974:251) refer to as the "importance of knowing how far modern attitudes are translated into modern actions," which is the primary microsocial link in the psychological modernity thesis of national development.

DATA AND METHODS

The setting for this investigation is Costa Rica, a small Central American re-

public with a democratic government, no army, extensive social security and health insurance schemes, and a literacy level of 89%. The largely agrarian population of 1.8 million had a GNP per capita of \$560 in 1970, which is 46th out of 122 countries compared by the World Bank (1972) and about equal to the average development level of nations involved in previous modernity research.

The data are from a 1971 survey of 210 Costa Rican adult males stratified by

rural/urban residence and by education level. Information was collected by means of structural interviews administered in Spanish by a trained team of six male Costa Rican university students. Sampling of respondents was conducted in several districts in the metropolitan area of San Jose and in seven rural areas 15 to 60 kilometers from the capital. Purposive quota sampling was used to select approximately equal numbers of respondents within each of four education levels in both rural and urban areas. After eliminating questionnaires with excessive missing data or response patterns, the final sample consisted of 28 respondents with no schooling, 58 with one or more years of primary schooling, 70 with some secondary schooling, 54 with some post-secondary schooling, and an equal number of rural and urban respondents in each category. Further details of data collection are reported in Armer and Schnaiberg (1975).

Psychological modernity was measured by Inkeles and Smith's (1974:348-50) Short Form 5 (OM-11) scale and scoring procedures are identical to those they report with dichotomization of items at the median. The omega unit-weighted maximum likelihood reliability (Heise and Bohrnstedt, 1970) is $\Omega_{\text{uwm}} = .633$, which is comparable to the Spearman-Brown coefficients (ranging from .62 to .74) for the OM-11 scale reported by Smith and Inkeles (1966:367).

Behavior measures

The survey included questions about fifteen behaviors which generally correspond to modern behavior discussed in the literature. For example, individuals who attain psychological modernity presumably

are more active in voluntary organizations and participate more in politics; they practice birth control more regularly and have fewer children as a result; they are quicker to adopt innovative practices in agriculture and are more productive as workers in industry . . . and, in general, they press more actively for social changes. (Inkeles, 1976:129)

Behaviors measured for respondents in the present study which approximate

those listed here are voluntary association membership (1); discussion of politics with friends (2); use of birth control methods (3); number of children (4); age at marriage (5); and productive employment vs. unemployment (6). Inkeles and Smith (1974:27) have also stated: "Religion ranks with the extended family as the institution most often identified both as an obstacle to economic development and as a victim of the same process." Data were available in our study on two behaviors that reflect religious and familial involvements; namely, frequency of prayer (7) and frequency of contact with family of origin (8). In addition to these behaviors there are others explicitly included in modernism measures such as mass-media participation by Kahl (1968) and husband/wife interaction patterns by Schnaiberg (1970). In the present study, indicators are available to measure use of radio, newspaper and cinema (9-11); support for wife's participation in the labor force (12); and for other women's rights, such as allowing wife to join in conversation with men visiting one's home (13), and visiting with men or women one doesn't know (14-15). (A list of behavior items and distributions is available on request.)

This catalogue of fifteen behaviors is neither comprehensive nor does it include behavior potentially detrimental to development such as excessive luxury-item consumerism, ethnic discrimination, and support of extended family members. Although the set of behaviors included in the original data was not guided by the present research objectives, it is a reasonably good sampling of behaviors suggested in the literature which, when cumulated across large numbers of individuals, "may become a collective input essential to the overall success of any program of national development" (Inkeles, 1976:129).

To provide an estimate of the general magnitude of the psychological modernity effect, we constructed a composite index of modern behavior. The index included the eight dependent behavioral variables which apply to the total sample rather than just to the married subsample, minus two items with lowest item-to-scale corre-

lations (family contact and prayer frequency) which contributed disproportionately to low scale reliability. Scale construction procedures repeat those used for the psychological modernity scale. The resulting six-item behavior index had a mean of .56, standard deviation of .25, and an omega coefficient of $\Omega_{uwml} = .533$ which was not significantly improved by eliminating further items from the scale. This modest internal consistency reliability will be taken into account when interpreting the net effects of background determinants and psychological modernity on the behavior index.

The background variables which might directly account for these modern behaviors and render psychological modernity effects spurious include age, urban residence, years of education, occupational status, and income level (see Table 1, for means, standard deviations, and scoring procedures). Religion, ethnicity, race, factory/nonfactory experience, and sex are wholly or largely invariant in our sample. Although researchers have investigated additional background variables as determinants of modernity, those mentioned here are the most common and influential.

Analysis Procedures

Data analysis consists of a series of ordinary least-squares (OLS) regression equations that assess the independent and the mediating effects of psychological modernity on the various indicators of modern behavior. We estimated two equations for each dependent variable—one including only the background variables and the second including those variables plus the measure of psychological modernity (OM-11). Entering psychological modernity into the equation permits assessing (1) the magnitude of its direct effect on the dependent behavioral variable, and (2) the extent to which it mediates the effects of background variables. We also report R^2 values in order to assess the ability of psychological modernity to increase explained variance in the dependent variable over the explanatory power of the background variables.

In each test we examine the effect of the

modernity variable and the reduction of the effects of other background variables when modernity is added to the equation. For each test there are four general outcomes possible. (1) If background factors do not channel their effects through psychological modernity but modernity independently influences behavior, then we should find a large effect of psychological modernity on the behavioral indicator (with a substantial increment in R^2) and no major decline in the effects of the background variables. (2) If modernity is an important mediator but not a independent determinant, then we should find a direct effect of modernity (and no increment in R^2 but a substantial decline in the independent effects of one or more background variables. (3) If modernity is relatively independent as a determinant of behavior *and* also mediates background effects, then we should find both a strong direct effect and a substantial reduction in effects from background variables. (4) Finally, if modernity is largely irrelevant (e.g., psychological modernity and behavioral measures are uncorrelated or spuriously related), then we should find neither a substantial modernity effect nor a major decline in the independent effects of background variables.

In testing the effects of psychological modernity and background variables, we used only cases for which complete data were available on all six independent variables ($N = 210$). Sample size is occasionally less when data were missing on the dependent variables but N 's are the same within sets. Tests of statistical significance are not technically warranted, especially in the several cases of dichotomous dependent variables, but we include them in order to indicate whether observed effects are haphazard or systematic (Winch and Campbell, 1969). We base our interpretations of substantially important effects on standardized regression coefficients greater than or equal to .14.¹ This decision point includes some

¹ Statistical difficulties are associated with regression equations involving binary dependent variables; but the estimates of the β 's remain unbiased (see Goldberger, 1964; or Kmenta, 1971). We rest our argument on the standardized regression coefficients.

coefficients which are not statistically significant but acknowledges any case in which psychological modernity accounts for as little as 2% of the explained variance in behavior. Although this decision point may favor confirmation of the modernity thesis, the results are such that selecting somewhat higher or lower decision points would not change the general conclusions of the study.

ANALYSIS

Analysis of Specific Behaviors

The first behavior tested in Table 1 shows that the background variables alone collectively account for 33% of the variance in movie attendance. Age, residence, and years of formal education all have substantial positive effects on attendance. When psychological modernity is added to the equation (second line in the set), it has only a small, negative coefficient ($\beta = -.029$) and the variance explained in membership ($R^2 = .334$) remains unchanged. Also, psychological modernity does not reduce, to an appreciable extent, the direct effects of any of the background variables. Overall, psychological modernity is largely irrelevant in accounting for movie attendance. The same pattern of results holds for the second behavior variable, radio listening, except that none of the background variables shows a sizable effect, and they jointly explain only 2.5% of the variance.

In general, the results for all fifteen dependent variables show that for four behavior variables psychological modernity has a sizable independent effect and mediates the effects of background variables (mainly education). (One of the four is not statistically significant at a .05 level.) For the other eleven dependent behavioral variables, the effect of psychological modernity does not exceed .12 (and is theoretically reversed in over half of the cases), and it does not mediate the influences of background variables to any substantial degree. Psychological modernity adds less than 2% to the explained variance (R^2) for twelve of the fifteen dependent variables, and no more than 3.5% for the remaining three.

The four behavioral variables for which

Table 1. Effects of Background Variables and Psychological Modernity on Behaviors of Costa Rican Adult Males

Dependent Variable	Background Variables				Psychological Modernity	R ²	N
	Age	Residence	Education	Occupation			
	Standardized Regression Coefficients						
(1) Attend Movies	.368*	.183*	.271*	-.111	—	.334	208
	.371*	.182*	.288*	-.110	-.029	.334	
(2) Listen to Radio Often	.002	.025	.053	-.074	—	.025	209
	.000	.026	.037	-.075	.029	.025	
(3) Read Newspapers Often	-.015	.052	.522*	.120*	—	.441	207
	-.026	.059	.424*	.113	.170*	.455	
(4) Organization Memberships	.020	-.065	.181*	.242*	—	.177	210
	.022	-.066	.191*	.243*	-.018	.177	
(5) Discuss Politics	.123	.007	.309	-.006	—	.168	208
	.105	.018	.153	-.017	.272*	.203	
(6) Employed/Unemployed	-.056	-.122	-.349*	.393*	—	.205	207
	-.061	-.119	-.401*	.387*	.079	.208	

Table 1. Continued

(7) Less Family Contact	-.023	-.050	-.302*	-.025	-.123	—	.159	209
	-.020	-.052	-.270*	-.023	-.112	-.056	.161	
(8) Less Frequent Prayer	.082	.194*	.185*	.020	.039	—	.111	207
	.064	.205*	.028	.008	.009	.274*	.146	
(9) Age at Marriage	-.234*	-.016	-.199	.273*	-.052	—	.119	99†
	-.234*	-.017	-.187	.273*	-.044	-.028	.120	
(10) Fewer No. of Children	.307*	.041	.513*	-.195	-.030	—	.323	99†
	.307*	.041	.488	-.199	-.048	.056	.329	
(11) Ever Use Contraceptives	.003	.089	.222	.063	.067	—	.105	83††
	-.010	.088	.127	.050	.012	.200	.126	
<i>Wife is Not Forbidden to:</i>								
(12) Take a Job Outside Home	.058	.131	.004	.273*	-.067	—	.094	99†
	.058	.131	.011	.274*	-.062	-.014	.094	
(13) Join in Visits with Men	.033	-.150	.142	.043	.203	—	.141	97†
	.034	-.149	.099	.035	.172	.098	.145	
(14) Talk to Men You Don't Know	.163	-.008	.104	.112	-.078	—	.062	98†
	.164	-.006	.051	.102	-.116	.120	.068	
(15) Visit Women You Don't Know	.072	.002	.091	-.075	.049	—	.019	98†
	.071	.000	.139	.067	.084	-.110	.025	
Mean	31.60	1.50	7.90	28.40	2.20	.66		
Standard Deviation	10.40	.50	5.10	25.30	.69	.22		

Note: All background variables are ordered in the direction of theoretically predicted positive association with modernity, e.g., old age to young, rural to urban residence, etc. Variable identifications are Age: self-reported years; Residence: rural (small settlements 15 to 60 kilometers from San Jose)=1, urban (San Jose)=2; Education: self-reported years of schooling; Occupation: self-reported description coded to Duncan SEI scores; Income: self-reported income below \$300=1, \$300 to \$800=2, above \$800=3, where 1 colonies=12; Psychological Modernity: Inkeles and Smith's OM-11 scale based on mean scores (0.00 to 1.00) across ten dichotomous items coded traditional=0 and modern=1. Items and distributions are available on request.

* Statistical significance of $p \leq .05$.

† Married subsample (N=99).

†† Married, Catholic subsample (N=89).

psychological modernity shows a nonnegligible effect are (3) newspaper reading, (4) less frequent prayer, (6) political discussion, and (11) contraceptive use. These behaviors share no obvious substantive characteristics that set them apart from the other behaviors and would help account for the observed modernity effects. However reading of newspapers is sometimes treated as a determinant of modern attitudes rather than as a "behavioral manifestation of individual modernity" (Inkeles and Smith, 1974: 252, 283). Thus, in this case, the observed path coefficient of psychological modernity may indicate the influence of newspaper reading on modern attitudes, rather than vice versa (i.e., misspecification of the model).

The fourth behavior, contraceptive use, is also problematic. Given the Catholic cultural setting of study, birth control behavior would be nonnormative for most respondents, and answers to the question on use of contraceptives are likely to be restrained and unreliable. In view of the high rate of nonresponse to this item alone (10%), the observed effect coefficients must be regarded as tentative.

The remaining two behaviors—frequent discussion of political issues and infrequent prayer—are not subject to these difficulties. Psychological modernity, or some subtheme incorporated within the psychological modernity scale, appears to have a modest but definite effect on these behaviors. One possible subtheme which might be responsible for several of the behavioral effects, especially infrequent prayer and contraceptive use, is religiosity or sacred vs. secular orientation to life.

We have made a preliminary investigation of this explanation for the four cases in which psychological modernity shows an effect on behaviors of individuals. Evidence of the importance of religiosity for these behaviors is observed when the single religious attitudinal item in the modernity scale is withdrawn and combined with an additional religiosity item in the survey to form a separate religiosity index.² When the religiosity index is in-

cluded in the regression analyses of the four behaviors, the effect of psychological modernity is reduced substantially (below statistical significance levels) for three—.128 for newspaper reading, .131 for prayer frequency, and .129 for contraceptive use. Psychological modernity continues to have a strong effect only on frequency of political discussion (.331). The religiosity index itself has a strong positive influence on prayer (.363) and contraceptive use (.302), and a negligible effect on newspaper reading and political discussion.

The general implication of these results is that the overall psychological modernity scale may include within it a subscale such as religiosity which can account for the previously observed behavioral effects of psychological modernity. Insofar as psychological modernity is conceived as important for behaviors relevant to national development such as those tested here, the influence might be better understood as the result of religiosity or some other as yet unspecified but well-known and less complex variable.

In sum, for most of the modern behaviors, psychological modernity has little or no independent or mediating influences. For the several behaviors that do not conform to the general pattern, psychological modernity effects are clear, though not large. We are unable to explain why psychological modernity is apparently contributing to these few behaviors and not to the others, but we raise the possibility that the effects are due to sacred-secular orientation, or some other subtheme incorporated in the modernity syndrome.

OLS Analysis of Behavior Index

An estimate of the general magnitude of the modernity effect and background variable effects on modern behavior is reported in Table 2. The first line shows the regression of psychological modernity on

² The OM-11 religiosity item is: "Do you think a man can be truly good without having any religion at all? Yes, No." The other religiosity item available in

the data is: "A man's wife is gravely ill. He obtains the best medical care and also prays fervently. She finally recovers. Which do you think was important in her recovery? Prayer alone, Mostly prayer, Mostly medical care, Medical care alone."

Table 2. Effects of Background Variables, Psychological Modernity, and Behavior Index

Dependent Variable	Background Variables					Mod.	R ²	N
	Age	Res.	Educ.	Occ.	Inc.			
Psychological Modernity (OM-11)	.066	-.042	.575*	.042	.182*		.525	210
Behavior Index	.163*	-.003	.438*	.058	.176*		.405	210
	.154*	.003	.359*	.052	.151*	.138	.414	210

* significance: $p \leq .05$.

the background variables. Education has the greatest effect on psychological modernity, a finding also reported by previous research (Kahl, 1968; Inkeles, 1969; Portes, 1973; Inkeles and Smith, 1974; Inkeles and Holsinger, 1974). Income also has a sizable net effect. None of the other coefficients exceed .14, nor depart significantly from zero. The five background variables account for 52.5% of the variance in psychological modernity.

The second line of Table 2 shows that education has the greatest effect on the behavior index as well, and that age and income have net effects that exceed .14 (and are statistically significant). When psychological modernity is added to the equation (line 3), these background effects are not substantially reduced. The direct effect of psychological modernity is only .138 (and statistically insignificant). This represents less than 2% of the explained variance in modern behavior and a negligible increase of 0.9% over variance explained by the background variables.

The implication of these results for the modernity thesis is clearer if the separate and shared effects of the background variables are combined and estimated for the path model shown in Figure 2. The direct effect of psychological modernity on modern behavior is positive but modest in

size (.138), while the background factors show strong effects on both the psychological modernity (.725) and behavior index (.536). Decomposing the total background effects (.636) into direct (.536) and indirect (.100) components shows the limited role psychological modernity plays in mediating background influences on behavior. These results support the alternative model of the relationship among background, modernity and behavior (Figure 1,B) rather than the traditional model (1,A) prevalent in modernity research. Psychological modernity as measured here plays a relatively insignificant role in explaining modern behavior.

Corrected Measurement Analysis of Behavior Index

While these summary results tend to run counter to the modernity thesis, the scale reliabilities of both psychological modernity and behavior are quite modest, as noted above, and may be attenuating the observed modernity effect. Also, the background variables are not perfectly measured and we can form a socioeconomic status composite of the education, occupation and income items in order to estimate and to correct for unreliability in background status. The SES composite

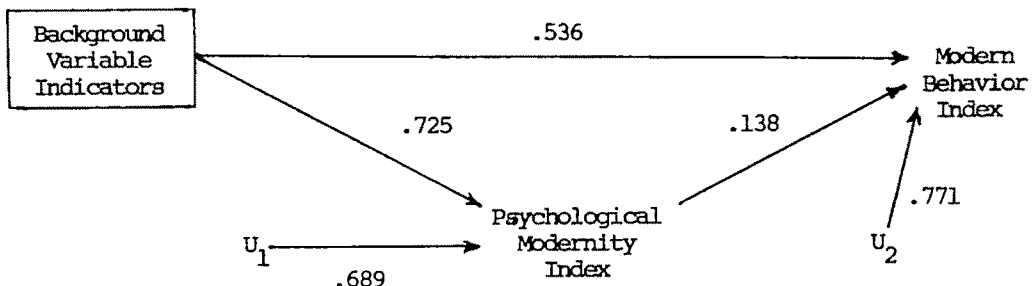


Figure 2. Estimated Model of Background Variables, Psychological Modernity and Modern Behavior

was formed by trichotomizing the ranges of the three status variables and summing across the item scores for each individual. The omega reliability is $\Omega_{uwm1} = .683$.

Our intention was to perform a summary two-equation regression analysis on the matrix of corrected correlations among age, socioeconomic status, psychological modernity and behavior: (1) Modernity = $f(\text{SES}, \text{Age})$; (2) Behavior = $f(\text{SES}, \text{Age}, \text{Modernity})$. However, the regression results are uninterpretable. Two standardized regression coefficients exceeded unity—the effect of SES on behavior and the effect of modernity on behavior—and the latter effect of modernity on behavior was negative.

These results were due largely to the extremely high intercorrelations between SES, modernity and behavior after correcting for attenuation.³ Socioeconomic status and psychological modernity have a corrected correlation which is close to unity (.977) and they apparently measure a common phenomenon. The true score variance on the psychological modernity index apparently estimates socioeconomic status or status-related attitudes. True scores on the modern behavior index also appear to be estimates of status (corrected $r = .927$). The most plausible causal interpretation of these results (if the modernity indices are not simply surrogate indicators of SES) is that socioeconomic status causes the status-related attitudes (measured by the modernity index). The status attitudes may, in turn, produce status-related behavior (measured by the modern behavior index) or the behavior may result directly from socioeconomic status. The high intercorrelations preclude a direct analysis of the magnitude of these alternative paths and lead us to consider compromise analysis strategies.

³ Observed correlations are corrected using omega unit-weighted maximum likelihood reliabilities, and the following correction formula:

$$r_{ij} = \frac{r_{ij}}{\sqrt{r_{ii}} \sqrt{r_{jj}}} \text{ (see, Bohrnstedt and Carter, 1971).}$$

High corrected correlations result for SES-Psychological Modernity (.977), SES-Behavioral Modernity (.927) and Psychological and Behavioral Modernity (.899). Age has modest correlations with SES (-.177), Psychological Modernity (-.284), and Behavioral Modernity (-.429).

FIML Analysis of Specific Behaviors

One compromise is to eliminate the unreliable behavior index and deal with specific behaviors as in the original OLS analysis. We originally treated the OLS behavior equations as unrelated—which may or may not be warranted. If they are in fact unrelated, then OLS analysis applied seriatum yields best, linear, unbiased estimates. However, it is quite likely that at least some of the behaviors, while not linked causally in theory, may in fact be linked via their error structure. An alternative specification of the model which takes this possibility into account involves what econometricians have referred to as “seemingly unrelated regressions” (Kmenta and Gilbert, 1968; Kmenta, 1971:517–29).⁴

As a further compromise, we may specify a model that takes measurement error into account for either SES or modernity, but not both simultaneously because, as we observed, correcting for both produces a unitary measure and uninterpretable results. We have chosen to treat the background variables as perfectly measured and to incorporate the measurement structure of modernity into the model. Such a model will likely lead to overestimation of modernity effects relative to background effects and may be interpreted as an upperbound of possible modernity effects.⁵

As a result of these considerations, we have estimated a modified structural model which incorporates the fallible measurement of psychological modernity alone and treats the eight behavioral equations as an entire system (see Figure 3).⁶

⁴ If the equations are linked through their disturbances, OLS applied seriatum will yield unbiased and consistent estimates, but the efficiency of the estimates is questionable. The gain in efficiency over OLS of certain alternative estimation procedures (e.g., generalized least-squares or maximum likelihood) has been shown to be a positive function of the correlations among disturbances and an inverse function of the correlations among explanatory variables (Kmenta, 1971:524).

⁵ However, measurement error induced bias is especially troublesome in cases beyond the bivariate and the direction of bias is not necessarily downward (Bohrnstedt and Carter, 1971).

⁶ The model in Figure 3 is described by the following system of structural equations:

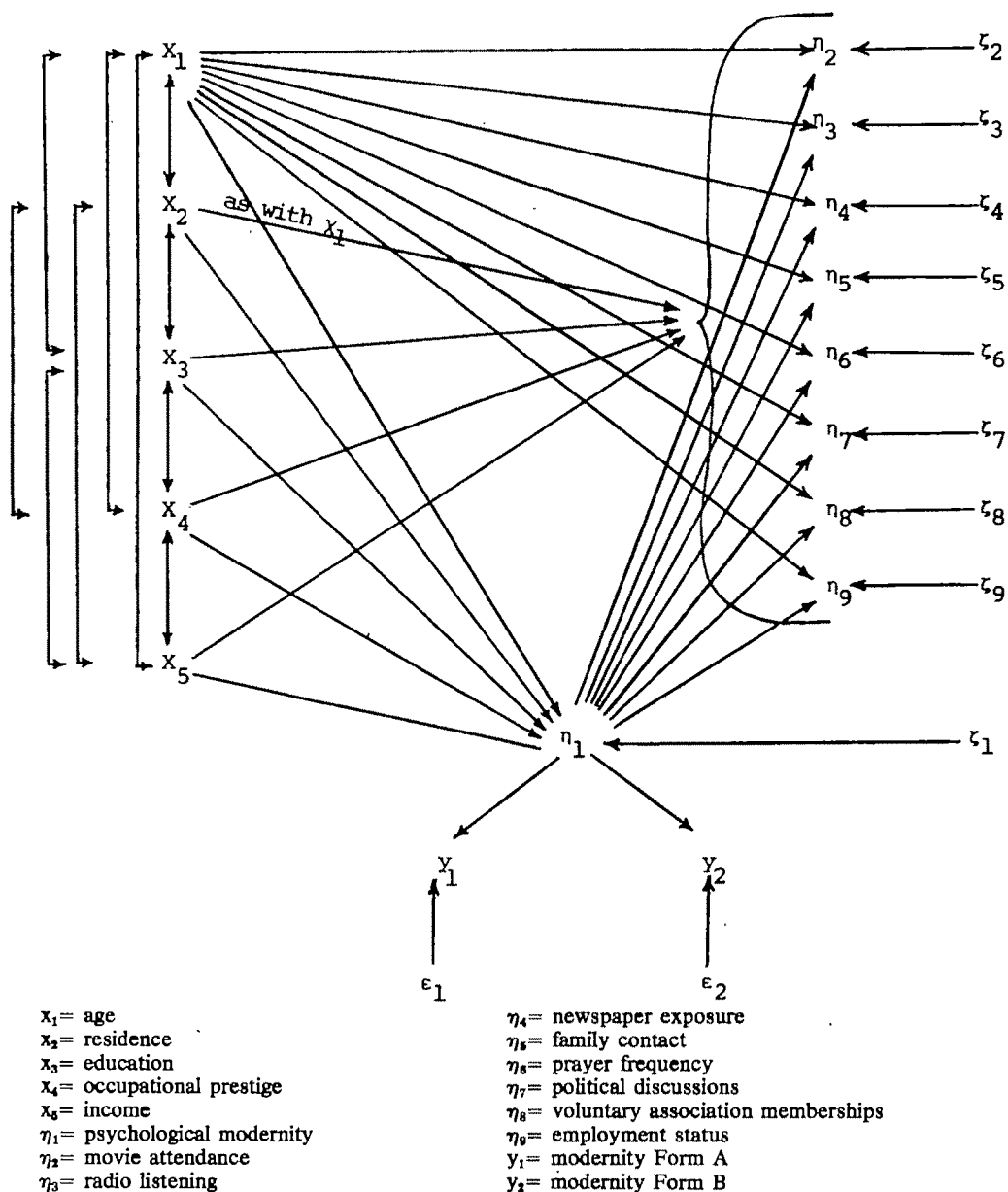


Figure 3. Seemingly Unrelated Regression Structural Model of the Eight Total Sample Behaviors

$$\eta_1 = \gamma_{11}x_1 + \gamma_{12}x_2 + \gamma_{13}x_3 + \gamma_{14}x_4 + \gamma_{15}x_5 + \zeta_1;$$

$$\eta_2 = \gamma_{21}x_1 + \gamma_{22}x_2 + \gamma_{23}x_3 + \gamma_{24}x_4 + \gamma_{25}x_5 + \beta_{21}\eta_1 + \zeta_2;$$

$$\eta_3 = \gamma_{31}x_1 + \gamma_{32}x_2 + \gamma_{33}x_3 + \gamma_{34}x_4 + \gamma_{35}x_5 + \beta_{31}\eta_1 + \zeta_3;$$

•

•

•

$$\eta_9 = \gamma_{91}x_1 + \gamma_{92}x_2 + \gamma_{93}x_3 + \gamma_{94}x_4 + \gamma_{95}x_5 + \beta_{91}\eta_1 + \zeta_9;$$

and the fallible measurement structure is given by:

$$y_1 = \lambda_{11}\eta_1 + \epsilon_1;$$

$$y_2 = \lambda_{21}\eta_1 + \epsilon_2.$$

We estimated the model with a full-information maximum likelihood (FIML) program (Jöreskog and Sörbom, 1976). This program is well-suited for present purposes since it allows a simultaneous estimation of all model parameters including correlated disturbances and the measurement structure. Additionally, the program provides a likelihood chi-square goodness-of-fit test and matrices of first-order partial derivatives of the fitting function (F) with respect to the individual

parameters of the model. This last feature produces information which can be useful in respecifying the initial model; for example, freeing-up restricted parameters (Sörbom, 1975).⁷ In the present application we were especially interested in determining whether the disturbance variance/covariance matrix for behaviors was in fact diagonal.

Figure 4 shows the results of the final estimated model containing the eight behaviors relevant to the total sample. Five parameters in the model are fixed rather than estimated in the total structure, that is, $r_{23} = .000$ due to the stratified sampling design, and the epistemic parameters for modernity were determined a priori by internal consistency reliabilities.⁸

Parameter estimation began by estimating a structure which corresponds to Figure 3, that is, with correlations among disturbances restricted to zero. The final parameter estimates in Figure 4 were arrived at by sequentially relaxing the recursiveness assumption concerning the disturbances by freeing-up the individual restricted parameter corresponding to the

largest first-order partial derivative of largest absolute value.

The final model presented in Figure 4 shows a reasonably good fit ($\chi^2_{40} = 50.012$, $p = .1333$) to the data. At this point 80 parameters are estimated for the model, which is the upper limit for the program. Compared to the earlier OLS results, the overall pattern of effects remains largely unchanged. Both income and especially education remain the major determinants of psychological modernity, and the effects are now larger in both cases. However, the previously observed direct effects of education on the behaviors have been washed out. The effect of education on the eight behaviors is entirely indirect via alteration in the attitudinal complex of psychological modernity.

Psychological modernity shows sizable significant effects on four of the eight behaviors—(η_5) family contact, (η_6) prayer frequency, and (η_7) political discussion—all but the effect on family contact showed up in the previous OLS results. However, the modernity effect on family contact is negative, opposite the direction predicted by modernity theory. The remaining half of the eight behaviors are not significantly affected by psychological modernity.

In Table 3, the explained variance of the eight individual behaviors is decomposed into three sources—that due to background variables exclusive of modernity (column 1), background variables and modernity jointly (column 2), and modernity alone (column 3). The first three columns sum to the total variance explained (R^2) in column 4. (Columns 5 and 6 are discussed later.)

⁷ More specifically, the fitting function referred to is given by the following expression:

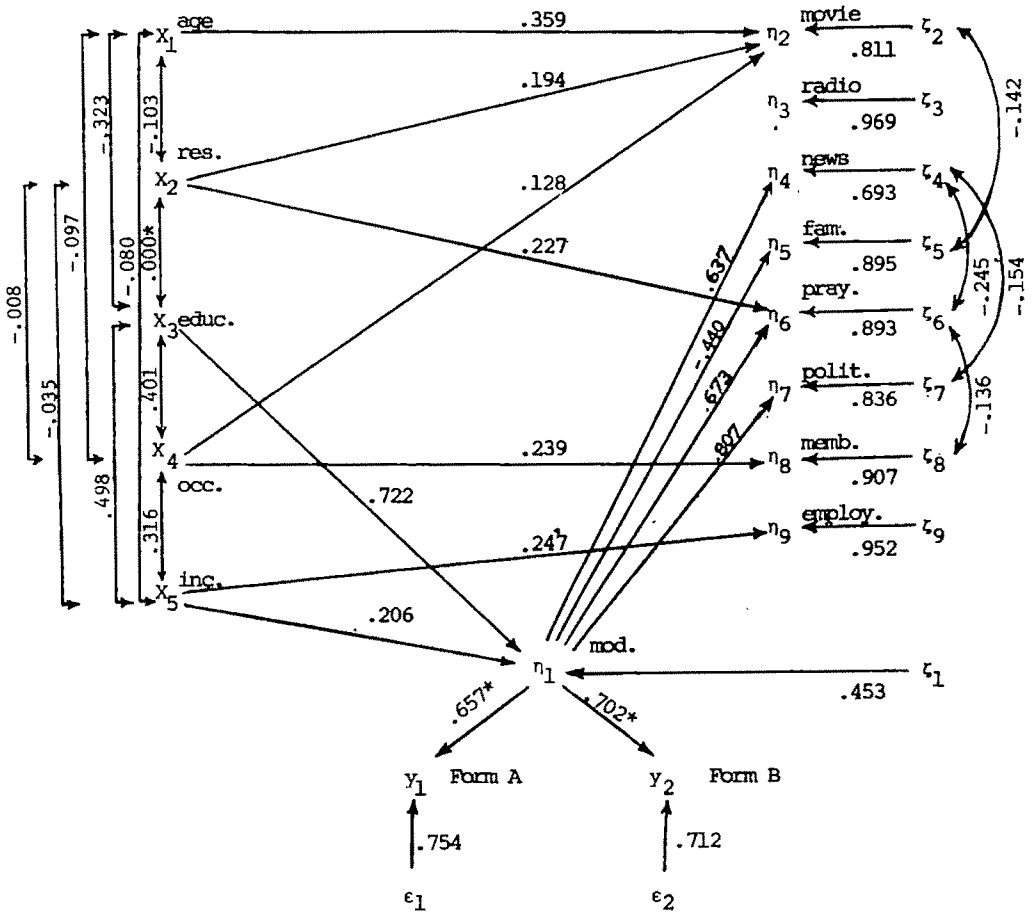
$$F = \ln |\Sigma| + \text{tr} (\Sigma \Sigma^{-1}) - \ln |S| - (p + q),$$

and is minimized with respect to θ , the vector of all independent, free and constrained parameters in the model (see Jöreskog and Sörbom, 1976). The procedure consists of examining the matrices, $\partial F / \partial \lambda_{ij}$, $\partial F / \partial \gamma_{ij}$, $\partial F / \partial \beta_{ij}$, $\partial F / \partial \gamma_{ij}$, $\partial F / \partial \phi_{ij}$, $\partial F / \partial \psi_{ij}$, $\partial F / \partial \omega_{ij}$ for the largest element. Then, the restricted parameter corresponding to the largest partial derivative of the largest absolute value is allowed to be estimated, assuming of course, that it is identified and makes theoretical sense to do so (see Sörbom, 1975 for further discussion).

⁸ The ten-item Inkeles OM-11 scale was randomly split into two five-item subscales. (This procedure was considered to be a reasonable compromise between employing the scale itself as a single-item indicator of modernity or allowing all ten items to be indicators of the construct. This latter option was precluded by the program capacity which allows for a maximum of 15 observed endogenous variables.) Then the omega reliability coefficient was computed based on the results of a confirmatory maximum likelihood factor analysis of all ten items together. The coefficients were $\Omega = .432$, and $\Omega = .493$ for Forms A and B, respectively. The square roots of these values yield the λ 's in the model and the error variances are readily computed by subtracting the reliabilities from 1.

⁹ The estimates given in Figure 4 represent four freeing-up estimations past the original recursive model. Each disturbance covariance which was estimated significantly improved the fit of the model to the data. The sequence of freed-up parameters and corresponding goodness-of-fit criteria were as follows:

Model	χ^2	df	DFR= χ^2/df	p
(1) Recursive	86.163	44	1.96	.0001
(2) Ψ_{64} , free	69.540	43	1.62	.0064
(3) Ψ_{64} , Ψ_{32} , free	61.929	42	1.48	.0243
(4) Ψ_{64} , Ψ_{32} , Ψ_{74} , free	54.853	41	1.34	.0726
(5) Ψ_{64} , Ψ_{32} , Ψ_{74} , Ψ_{88} , free	50.012	40	1.25	.1333



*Fixed parameters.

$\chi^2 = 50.012$, $df=40$, $p=.1333$

Note: See Figure 3 for variable names. All exogenous to endogenous effects are estimated. Only coefficients which are at least twice their standard errors are presented. Correlation matrix available on request.

Figure 4. Seemingly Unrelated Structural Equation Model of Behaviors with Psychological Modernity as an Unmeasured Construct (N=210)

For three behaviors (movie attendance, organization membership, and employment), the explained variance due to background variables alone substantially exceeds that due to modernity alone or in joint combination with background variables. In the case of radio listening behavior, the influence of background factors as compared to modernity is approximately equal and the total variance explained by both is quite low. For four behaviors (newspaper reading, family contact, prayer, and political discussion), the variance explained by modernity exceeds that explained by the background variables but the influence on family contact is opposite from the theoretical prediction. The explained variance in two of these

behaviors (prayer and political discussion) is actually reduced by joining background/modernity influences. The precise substantive meaning of this joint influence is not clear, but the reduction results from negative direct effects (statistically nonsignificant) of education, occupation and income when modernity is controlled.

The fifth column of Table 3 gives the reduced form R^2 , or the variance explained by the five background variables before modernity is added to the equations. The final column gives the difference between the full and reduced form R^2 's, or the incremental increase to explained variance due to modernity.

The increments in column 6 are all quite

Table 3. Behavior Variance Explained by Background Variables and Psychological Modernity from the Structural Model*

	Background			Explained Variance		
	Background Exclusive of Modernity (1)	Background and Modernity Jointly (2)	Modernity Exclusive of Background (3)	Total R ² (4)	Reduced Form ^a R ² (5)	Increment Due to Modernity ^b (6)
Attend Movies	.229	.062	.052	.343	.335	.008
Listen to Radio Often	.014	-.002	.016	.028	.025	.003
Read Newspaper Often	.022	.092	.406	.520	.439	.081
Less Family Contact	.004	.001	.194	.199	.160	.039
Less Frequent Prayer	.181	-.436	.457	.202	.111	.091
Discuss Politics	.121	-.471	.651	.301	.168	.133
Organization Memberships	.150	.025	.002	.177	.177	.000
Employed/Unemployed	.094	.000	.000	.094	.088	.006

* Based on the model presented in Figure 3. Results were calculated using all estimates, that is, even those which were not twice their standard errors and therefore do not appear in Figure 3.

^a The calculated reduced form R² gives the variance explained in each individual behavior due to the background variables alone (i.e., without psychological modernity being partialled out).

^b Gives the contribution to explained variance due to modernity beyond that explained by the background variables above.

small and exceed 5% in only three cases—newspaper reading, prayer, and political discussion. These are the same three behaviors which showed sizable modernity effects in the OLS analysis. The increase in explained variance due to modernity amounts to approximately 45% of the total explained variance (column 4) in prayer frequency and political discussion, and less than 20% in the remaining six behaviors.

While the model fits the data reasonably well, as mentioned above, the amount of variance explained in these eight behaviors is not high. The percentage of variance explained ranged from a high of 52% for newspaper reading to a low of 2.8% for radio listening. The fact that four of the behavior variable disturbances were significantly correlated suggests the existence of one or more common omitted determinants of the behaviors. The role of psychological modernity suggested by the estimated model is primarily as a channel of background effects on half of the behaviors, as an independent determinant of 8 to 13% of the variance in three of these behaviors, and as irrelevant for the remaining half of the behaviors.¹⁰

DISCUSSION

We began by raising questions concerning the role of psychological modernity as expressed in the modernity literature in determining behaviors considered to be important for national development efforts. Several types of analysis have provided mixed results, but at the same time several important consistencies remain. First, when fifteen behaviors are examined as separate equations, we find that only three (frequency of newspaper reading, prayer frequency, and involvement in political discussions) of the fifteen behaviors were affected significantly by psychological modernity and one other behavior (contraceptive use) showed an effect of psychological modernity which was not statistically significant, but sizable enough to be of substantive interest. Furthermore, comparing the R² of reduced-form equations with that of the structural form which includes psychological modernity demonstrated that psychological modernity contributes little to explained variance in most behaviors.

Second, when we turned to a summary composite behavioral index, we found

¹⁰ A similar analysis was carried out for the seven behaviors relevant to the married subsample only. In general the results were similar to the OLS counter-

parts in that modernity did not show a significant effect on any of the seven behaviors although some coefficients were sizable.

that (a) the set of eight total sample behaviors did not scale well (a subset of six behaviors with improved but continuing low reliability had to be used); (b) regression results using the behavioral and psychological modernity composites indicated a minor role for modernity in determining the general index of behavior; and (c) corrected correlations among an SES composite, age, modernity and behavior measures showed both psychological modernity and modern behavior to be almost perfectly associated with socioeconomic status—raising questions concerning the label of modernity, if not the meaning of the general constructs.

Third, our estimation of a full structural model of eight behaviors incorporating the measurement error of psychological modernity into the structure and examining the results for interrelationships among the dependent behaviors provides evidence generally consistent with the earlier separate equation findings. Four of the eight behaviors (newspaper reading, contact with family, frequency of prayer, and political discussions) are significantly influenced by psychological modernity. One of the four, family contact, did not show up in the separate equation results and is affected in the direction opposite that predicted by modernity theory. Also, psychological modernity explains little of the variance in most of the behaviors beyond that explained by background variables. In comparison with the earlier results, the effects on the four behaviors were consistently larger and the direct effects of education on behaviors were found to be absent. Education's influence appears totally indirect through the psychological measure.

The general pattern resulting from the several types of analysis is that psychological modernity (a) is important in determining only a limited number of behaviors examined, and (b) generally adds little to explaining behavioral variations beyond objective background characteristics. Rather, it appears more as an interpretative construct (and that only in several cases) than as a pervasive source of modern behaviors.

Overall the findings neither confirm the modernity thesis, nor do they consistently

disconfirm it. Rather, the most liberal interpretation is that psychological modernity may be an important channel or way of interpreting the influence of education and other background determinants on some modern behaviors. On the more conservative side, the results suggest that psychological modernity as measured here may be mainly a measure of social class values. A general conclusion crosscutting both interpretations is that modernity does not have the predicted effect on a majority of the behaviors. Even the FIML analysis of a model slanted in favor of modernity effects found direct influences for only half the behaviors, and one was affected in the opposite direction from modernity predictions. Whether one takes a more liberal or conservative stance with respect to interpreting the influence of psychological modernity in the foregoing analysis, the results undermine confidence in the modernity thesis regarding the importance of modern attitudes and values in producing behavior identified as instrumental for modernization of society.

There are several factors which may contribute to the pattern of these results which should be noted. First, the use of cross-sectional data and self-report measures precludes unambiguous causal analysis and raises questions about possible contamination and unreliable measurement. The possibility of contamination of modern attitudes and behaviors with each other or with invalidity factors as a result of cross-sectional data, self-reported items, or other measurement features would tend to overestimate the effect of psychological modernity. Likewise, our decision to take into account low reliability of the attitudinal measure while treating background variables as perfectly measured in the FIML reestimation contributes to the likelihood of confirming the modernity thesis.

Another factor which may account for partial support of the modernity thesis is the content of the modernity scale used in the analysis. Previous research (e.g., Armer and Schnaiberg, 1977) has suggested that this particular modernity scale has a high degree of socioeconomic status invalidity. Hence, the likelihood of

the effect of education being attenuated when psychological modernity is treated as intervening is probably higher with the OM-11 measure than if other modernity scales less confounded with SES had been employed.¹¹

A potential objection to our analysis is that we misrepresent the thesis as predicting a causal relationship between modern attitudes and behaviors when, in fact, only a simple association is claimed for these two aspects of individual modernity. It may be argued that modernity theorists do not regard psychological modernity as causing behavioral modernity; instead, both modern attitudes and behaviors belong to a syndrome of highly correlated characteristics defining modern man, and both are products of education and other modernizing social forces. Thus, all that is claimed is that men who are modern in attitudes are also modern in behaviors and are important for societal modernization. As a result, all that can be tested is the predictive validity of psychological modernity insofar as it correlates positively with modern behaviors.

Such an objection would be difficult to accept on the grounds that it is inconsistent with causal statements in the modernity literature. A single quotation to illustrate the point is Inkeles and Smith's (1974:313) argument that:

. . . [T]he modern man showed himself to perform differently from the more traditional man in many realms of action having practical bearing on the processes of societal modernization. The modern man is quicker to adopt technical innovation, and more ready to implement birth-control measures; . . . and he permits his wife and daughter to leave home for more active participation in economic life. In these and a host of other ways, only some of which we have documented, the man who is more modern in attitude and value acts to support modern institutions and to facilitate the general modernization of society.¹²

¹¹ If modernity is interpreted as virtually equivalent to education or SES (as do researchers who use education or SES indicators of modernity), then the thesis that modernization of individuals in society leads to socioeconomic development is largely tautological.

¹² The data also do not support this interpretation. If modern attitudes and behaviors are components of

Finally, while the behaviors examined above are representative of those in the modernity literature, we do not consider these to be the only behaviors worthy of investigation. For example, our data omitted many important economic, social, and political behaviors (e.g., consumerism, work productivity, savings behavior, social discrimination and voting). Furthermore, our results are from a sample drawn in one Latin American society. Similar research should be carried out with different behavioral measures and in other Third World and developed countries.

It is also worth emphasizing that we have examined only a limited portion of the modernity thesis. The modern behavior-societal change relationship is not examined here and there exists very little systematic evidence regarding the impact of aggregated orientations or behaviors on national development. Hence, the role of attitudes and behaviors in the development process is still open to question. In short, we do not deny that attitudes and behaviors of individuals or of aggregates may play an important role in promoting (or resisting) social change. Rather, our evidence suggests that psychological modernity has, at best, limited influence on social change through its impact on behavior.

The further argument that psychological modernity is important because it has a *direct* influence on the modernization of societies, net of its influence on behavior, remains unexamined. We note that this claim has been contradicted by others (e.g., Carnoy, 1974; Galtung, 1971: 93-4) who contend, for example, that psychological modernization is a form of cultural imperialism leading to greater dependence and underdevelopment of societies. Neither the modernization nor opposite view has been substantiated. Our evidence suggests that if psychological modernity contributes to either the development or the underdevelopment of

a single construct rather than of separate, causally related constructs, then their correlation should be higher with each other than with other constructs. In fact, their correlation is lower than the correlation of each with education or SES, even after correcting for unreliability.

societies, it does so primarily as a channel of personal background effects on behavior and not as an independent source.

Those who use the concept of individual modernity in their research or theory should note (1) the lack of evidence regarding the significance of modern attitudes for either national development or underdevelopment, and (2) the present evidence that psychological modernity has limited relevance for personal behaviors identified in the modernity literature as important for national development. In sum, the modernity thesis is generally unsupported except insofar as psychological modernity acts as a mediator of background influences on certain behaviors. Research is needed examining additional behaviors in other contexts and focusing on specific psychological and behavioral constructs rather than on the loose cluster of variables which make up the modernity measures. In addition, research is needed which aggregates psychological and behavioral characteristics and examines their effects on macro-dimensions of economic and national development over time. The role of orientations and behaviors as forces, conditions, or obstacles to social change cannot be ignored by those who seek a comprehensive understanding of social change processes, especially socioeconomic development or underdevelopment. The welcome focus of world-systems analysis should not blind researchers to internal processes that are also vital to the destinies of societies.

REFERENCES

- Armer, Michael and Allan Schnaiberg
1975 "Individual modernity, alienation, and socioeconomic status: a replication in Costa Rica." *Studies in Comparative International Development* 10:35-47.
1977 "Reply to Cohen and Till." *American Sociological Review* 42:378-82.
- Bohrnstedt, George W. and T. Michael Carter
1971 "Robustness in regression analysis." Pp. 118-46 in H. L. Costner (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.
- Carnoy, Martin
1974 *Education as Cultural Imperialism*. New York: McKay.
- Chase-Dunn, Christopher
1975 "The effects of international economic dependence on development and inequality: a cross-national study." *American Sociological Review* 40:720-38.
- Chodak, Szymon
1973 *Societal Development: Five Approaches with Conclusions from Comparative Analysis*. New York: Oxford University Press.
- Cohen, Jere and Amnon Till
1977 "Another look at modernity scales: reanalysis of the convergent and discriminant validities of the Armer, Kahl, Smith and Inkeles, and Schnaiberg scales (comment on Armer and Schnaiberg, ASR June, 1972)." *American Sociological Review* 41:373-8.
- Coleman, James S.
1971 "Conflicting theories of societal change." *American Behavioral Scientist* 14:633-50.
- Feldman, Robert H.
1975 "The relationship of school, grade and sex to traditional-modern attitudes among Gusii students in Kenya." *Journal of Social Psychology* 96:135-6.
- Galtung, Johann
1971 "A structural theory of imperialism." *Journal of Peace Research* 2:81-117.
- Goldberger, Arthur
1964 *Econometric Theory*. New York: Wiley.
- Gough, Harrison G.
1976 "A measure of individual modernity." *Journal of Personality Assessment* 40:3-9.
1977 "Further validation of a measure of individual modernity." *Journal of Personality Assessment* 41:49-57.
- Hechter, Michael
1975 "Review essay of Wallerstein: the modern world-system; capitalist agriculture and the origins of the European world-economy in the sixteenth century." *Contemporary Sociology* 4:217-22.
- Heise, David R. and George W. Bohrnstedt
1970 "Validity, invalidity, and reliability." Pp. 104-29 in E. F. Borgatta and G. W. Bohrnstedt (eds.), *Sociological Methodology*. San Francisco: Jossey-Bass.
- Holsinger, Donald B.
1974 "The elementary school as modernizer: a Brazilian study." Pp. 22-46 in Alex Inkeles and Donald B. Holsinger (eds.), *Education and Individual Modernity in Developing Countries*. Leiden: Brill.
- Horowitz, Irving Louis
1970 "Personality and structural dimensions in comparative international development." *Social Science Quarterly* 51:494-513.
- Inkeles, Alex
1969 "Making men modern: on the causes and consequences of individual change in six developing countries." *American Journal of Sociology* 75:208-25.
1971 "Continuity and change in the interaction of the personal and the sociocultural systems." Pp. 265-81 in B. Bernard and A. Inkeles (eds.), *Stability and Social Change*. Boston: Little, Brown.
1975 "Emerging social structure in the world." *World Politics* 27:467-95.

- 1976 "Understanding and misunderstanding individual modernity." Pp. 103-30 in Lewis A. Coser and Otto N. Larsen (eds.), *The Uses of Controversy in Sociology*. New York: Free Press.
- Inkeles, Alex and Donald B. Holsinger (eds.)
[1972] *Education and Individual Modernity in 1974 Developing Countries*. Leiden: Brill. (Reprint of *International Journal of Comparative Sociology* 13.)
- Inkeles, Alex and David H. Smith
1974 *Becoming Modern: Individual Change in Six Developing Countries*. Cambridge, Ma.: Harvard University Press.
- Jöreskog, Karl G. and Dag Sörbom
1976 *LISREL III: Estimation of Linear Structural Equation Systems by Maximum Likelihood Methods: A Computer Program Manual*. Chicago: National Educational Resources.
- Kahl, Joseph A.
1968 *The Measurement of Modernism: A Study of Values in Brazil and Mexico*. Austin: University of Texas Press.
- Kmenta, Jan
1971 *Elements of Econometrics*. New York: Macmillan.
- Kmenta, Jan and Roy F. Gilbert
1968 "Small sample properties of alternative estimators of seemingly unrelated regressions." *Journal of the American Statistical Association* 63:1180-1200.
- Lerner, Daniel
1965 *The Passing of Traditional Society*. Glencoe: Free Press.
- Peshkin, Alan and Ronald Cohen
1967 "The values of modernization." *Journal of Developing Areas* 2:7-21.
- Portes, Alejandro
1973 "The factorial structure of modernity: empirical replications and a critique." *American Journal of Sociology* 79:15-44.
1976 "On the sociology of national development: theories and issues." *American Journal of Sociology* 82:55-85.
- Portes, Alejandro and Jose A. Cobas
1976 "Modernism and behavior: the predictive values of subjective versus objective factors in a developing country." *International Review of Modern Sociology* 6:253-66.
- Rutinson, Richard
1976 "The world-economy and the distribution of income within states: a cross-national study." *American Sociological Review* 41:638-59.
- Schnaiberg, Allan
1970 "Measuring modernism: theoretical and empirical explorations." *American Journal of Sociology* 76:339-425.
- Smith, David H. and Alex Inkeles
1966 "The OM scale: a comparative socio-psychological measure of individual modernity." *Sociometry* 29:353-77.
1975 "Individual modernizing experiences and psycho-social modernity: validation of OM-scales in six developing countries." *International Journal of Comparative Sociology* 16:155-73.
- Sörbom, Dag
1975 "Detection of correlated errors in longitudinal data." *British Journal of Mathematical and Statistical Psychology* 28:138-51.
- Suzman, Richard M.
1977 "The modernization of personality." Pp. 40-77 in Gordon J. DiRenzo (ed.), *We, the People: American Character and Social Change*. Westport: Greenwood Press.
- Wallerstein, Immanuel
1974 *The Modern World-System: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century*. New York: Academic Press.
1976 "Modernization: requiescat in pace." Pp. 131-5 in Lewis A. Coser and Otto N. Larsen (eds.), *The Uses of Controversy in Sociology*. New York: Free Press.
- Witch, Robert F. and Donald T. Campbell
1969 "Proof? no. evidence? yes. the significance of tests of significance." *American Sociologist* 4:140-3.
- World Bank (IBRD)
1972 *World Bank Atlas 1972*. Washington, D.C.: International Bank for Reconstruction and Development.

BEHAVIORISM ON VERSTEHEN AND ERKLÄREN*

JOHN D. BALDWIN AND JANICE I. BALDWIN

University of California, Santa Barbara

American Sociological Review 1978, Vol. 43 (June):335-347

Modern behaviorism has made significant progress in analyzing private behavior: cognition and emotion. The behavioral distinction between contingency-shaped and rule-governed behavior establishes a starting point for analyzing tacit (subjective, private) knowledge and explicit (objective, public) knowledge. The present paper extends this analysis to include the issues of verstehen and erklären. One subset of tacit knowledge is verstehen. Behavioral analysis (1) describes the special qualities of this type of tacit knowledge and (2) delineates the circumstances under which verstehen will occur and produce valid understanding. Modern behaviorism makes it possible to carry out a scientific analysis of the socialization experiences through which people learn to attach meaning to their thoughts, actions and emotions. Thus, modern behaviorism can generate an erklärte understanding of the topic usually reserved for verstehen. Neither tacit nor explicit knowledge can replace the other, nor can verstehen and erklären. Both are needed and natural components of human behavior. Since both are frequently intermeshed, rarely does one encounter a pure form of either.

The argument in favor of verstehen has a deeply rooted position in the sociological tradition, with such important advocates as Dilthey, Rickert, Weber, Cooley and Mead. It has become a classic tenet in sociology that erklären is an inadequate epistemology for human studies and that sociology must be founded on verstehen. However, when Abel (1948), Munch (1957) and Wax (1967) attempted to delineate a precise explanation of verstehen, it became obvious that a shared meaning was lacking. Recently, numerous writers from the cognitive, interactionist, Weberian, phenomenological and humanistic schools have defended various versions of verstehen; the behavioral viewpoint has not yet been heard.

Logical positivism and operationism are well known for having abolished feelings, intention, meaning and other subjective experiences from their realm of study. Methodological behaviorism has followed this route in its attempt to establish lawful patterns in observable behavior; and with respect to its goals, methodological behaviorism is a successful and flourishing enterprise in modern ethology and psychology.

Behaviorists are, however, well aware of the problems inherent in logical positivism and operationism; and during the past decade there has been significant progress toward a more comprehensive behavioral philosophy and empirical epistemology that incorporates private experience, feelings, choice, self-control and values within the domain of issues studied (Skinner, 1953; 1969; 1974; Bem, 1964; 1970; Homme, 1965; Bandura, 1969; Mahoney, 1974; Thoresen and Mahoney, 1974; and others). Skinner (1969) criticizes the positivistic strategy of defining subjective experiences outside the range of scientific methodology, pointing out that it fails to cope with phenomena that are too important to be denied. Subjective phenomena are real and must be included in any theory of human behavior or else the theory is inadequate. Branches of modern behaviorism, commonly referred to as radical behaviorism (Skinner, 1974) and social learning theory (Bandura, 1969; Mahoney, 1974) have emerged to cope with the subjective events not covered by methodological behaviorism. A task of modern behaviorism is to

face the problem of privacy. It may do so without abandoning the basic position of behaviorism. Science often talks about things it cannot see. . . . An adequate science of behavior must consider events taking place

* We wish to thank B. F. Skinner, Kyle Kercher, Donald Cressey, John F. Scott, and Connie Veldink for critical feedback on the manuscript.

within the skin of the organism, not as physiological mediators of behavior, but as part of behavior itself. (Skinner, 1969:227-8)

Modern behaviorism gives rise to an important new perspective for evaluating the role of *verstehen* and *erklären* in the social sciences. Our goal is to extend the work of modern behaviorism by making a behavioral analysis of *verstehen* and *erklären*, along with the strengths and limits of each method.

TACIT AND EXPLICIT KNOWLEDGE

Bridgman (1952; 1959), Polanyi (1959; 1960; 1966) and Skinner (1969) make an important distinction between two crucially different kinds of human knowledge. *Tacit* knowledge is personal, subjective knowledge based on direct experience with the social or nonsocial environment; and *explicit* knowledge is public, verbally encoded information that has been passed on via symbolic mediation. Although these two types of knowledge are more general than the common conception of *verstehen* and *erklären*, they provide a starting point for a behavioral analysis of these latter concepts.

Tacit and explicit knowledge are extreme types, poles of a continuum. First, we will examine the behavioral principles that Skinner adduces to explain these two extreme types. This analysis will demonstrate that mixtures of the two types are more common than the pure types. Then we will partition human knowledge into three nested sets. Tacit knowledge in the third set—the understanding of action and attached meanings—corresponds to the classical Weberian conception of *verstehen*. According to a behavioral analysis, explicit knowledge in any of the three sets can be attained through the process of *erklären*.

CONTINGENCY-SHAPED VS. RULE-GOVERNED BEHAVIOR

When a person learns a behavior from direct experience, without any rules or verbal guidance from others, the behavior is said to be contingency-shaped be-

havior.¹ The child who learns not to touch bees after having been stung while playing in the clover has learned contingency-shaped behavior. When a person learns a behavior by following rules which are explicitly or implicitly encoded in verbal statements, the behavior is said to be rule-governed. After hearing explicit rules ("Don't touch bees") or implicit rules ("Bees will hurt you if you touch them"), a second child may show rule-governed avoidance even though he has had no personal experience with touching bees or being stung. Both children have learned to avoid bees and have acquired knowledge about bees. The physical appearance of their avoidance behavior may be very similar, but their knowledge differs in several important respects.

The child who has been stung has tacit knowledge based on personal discovery and real life experience, rather than on education. Through Pavlovian conditioning (also called respondent or classical conditioning), this child has learned conditioned emotional responses that will be elicited the next time he encounters bees;² and these contribute an emotional element to the personal feelings typically associated with contingency-shaped behavior. Although the child may be able to put some of his feelings into words if asked about his sting, the learning took place independently of verbal mediation or rules. Tacit knowledge need not rely on verbal elements to be vivid and subjectively real. A very young child might not be able to verbally encode his experience

¹ In the behavioral literature, the use of the concept "contingency-shaped behavior" can have two meanings. In one sense, all learned behavior is determined by the contingencies of reinforcement; and rules are seen as one subset of the total contingencies. In the present contrast between tacit and explicit knowledge, Skinner tends to use contingency-shaped behavior to stand for all behavior that is not rule-governed. This functions for the purpose of analysis to heighten the contrast between tacit and explicit knowledge as extreme types, though it will be shown later that the extremes are less common than intermediate forms in most aspects of life.

² For details concerning Pavlovian conditioning of emotional responses, see Reynolds (1968) or Staats (1968).

or knowledge; but we can tell by his subsequent behavior around bees that he has learned to fear and avoid them. The label "tacit knowledge" is used to indicate that there is a subjective quality to contingency-shaped knowledge, and that explicit verbal accounts are not necessarily involved, though they may be invented on an ad hoc basis after the learning experience.

The child who has learned to avoid stings by following rules also has knowledge about bees, but his knowledge contains less personal feeling than tacit knowledge does. If the parents were relatively unemotional while telling their child not to touch the little yellow and black insects in the clover, the child would undergo little Pavlovian conditioning and have few conditioned emotional responses associated with bees. If an overanxious parent expressed high levels of emotionality or used anxiety arousing words while telling rules for avoiding bees, the child would acquire stronger emotional responses through this aversive conditioning. However, the emotional feelings were derived from social interaction rather than from direct personal experiences with bee stings. The child still does not know what a bee sting feels like, though he may fear it. His knowledge is abstract and his emotional responses are secondhand, with no guarantee that they accurately parallel the emotional responses conditioned by personal experience. When the child is asked to explain what he knows about bees, his verbal accounts will be highly influenced by the rules he learned. Because the rules were given to the child in a verbally encoded form, they lend themselves to explicit verbal output when he is asked about his knowledge. Rules are preformulated knowledge that can be quickly transposed into answers.

ERKLÄREN

One main drawback with contingency-shaped knowledge is that it cannot be passed from one person to the next.

By its very nature every embodied spirit is doomed to suffer and enjoy in solitude. Sensations, feelings, insights, fancies—all these

are private and, except through symbols and at secondhand, incommunicable. We can pool information about experiences, but never the experiences themselves. (Huxley, 1954:12)

Contingency-shaped knowledge is private knowledge: a felt, intuitive kind of understanding. If a person attempts to encode his tacit knowledge into a verbal message, he may succeed in passing *something* to a second person; but the verbal information that is passed will have the properties of a rule for the second person (assuming that the second person does not have prior contingency-shaped knowledge about the experiences being discussed).³ If the second person listens to the message and utilizes the encoded information, the second person's behavior will be rule-governed rather than contingency-shaped, and it will not be associated with the kinds of tacit knowledge—feelings, emotional responses and attached meanings—found in the first person's behavior.

Even though rules never capture all of the wisdom, feelings and sensitivity to individual cases that are found in tacit knowledge, people learn to be rule creators: the rewards of passing on information and helping others to learn rapidly condition people into becoming rule creators.⁴ A novice can be introduced into

³ If the second person has similar contingency-shaped (tacit) knowledge, the message will be more than a rule, since it may evoke memories and feelings of shared experiences in the second person. But the memories and feelings originate from the second person's own past conditioning and are not created de novo by the verbal message.

⁴ Most people learn how to ad hoc rules to describe their contingency-shaped behavior. Therefore, hearing a person give an account of his behavior in terms of rules is not evidence that the behavior being described was actually controlled by those rules. If a sociologist takes these rule descriptions literally and later assumes the person's behavior was controlled by a rule or norm, the descriptions will be misleading to the degree that the behavior was actually contingency-shaped.

Since much of human behavior is contingency-shaped, it is not surprising that the ethnomethodologists (e.g., Garfinkel, 1967:18–24) have found that most people cannot ever capture the complexity of their own behavior in a set of rules and hence need to ad hoc ever more rules to explain how they would use, bend, or supersede any other rules they may

a complex behavioral pattern or body of verbal knowledge very rapidly via rules; whereas it might take years to learn the behavior or knowledge by personal experience. A unique part of being human is that we can benefit from verbally encoded information. Our verbal culture is filled with maxims, rules of thumb, customs, proverbs, shop lore, aphorisms and other types of folk knowledge that aid in learning sports, crafts, manual skills, art, creative writing, philosophy, and so forth. Even though the information is second-hand and often crude, culturally accumulated rules help people learn more than any single person could learn by himself from personal experience with the contingencies.

Scientific knowledge is a collection of verbally (or mathematically) encoded rules that allow us to predict or control a portion of our environment (Skinner, 1969). Scientific knowledge differs from folk knowledge only in degree. The scientific method provides checks and tests which usually operate to create more pragmatically useful—and hence more reinforcing—rules than folk methods do. It is the utility of scientific rules that has reinforced us for relying more and more heavily on science over the past 400 years. Newtonian physics, for example, describes the behavior of many kinds of objects in space and time, and it contains the rules for predicting and controlling the behavior of those objects. Bridge builders, architects and scientists who utilize this rule-contained knowledge can learn more about the behavior of objects than any single person could ever learn by himself or from folk knowledge. "Objective knowledge transcends the individual; it is more stable and durable than private experience . . ." (Skinner, 1969:155). However, explicit knowledge has its origin in personal knowledge, the same as maxims and folk knowledge do. The main dif-

ference between science and folk knowledge is that the scientific community provides a system of public challenge, negotiation and falsification that puts stronger contingencies for careful observation, rule-formulation and testing on its practitioners than the nonscientific community does. "For science gets at the truth not so much by avoiding mistakes and personal bias as by displaying them in public—where they can be corrected" (Commoner, 1971:84).

When scientific rules fail to predict and control in some area of research, new rules may be devised to replace the prior set. The behavioral position is compatible with the view that there is no ultimate or absolute set of empirical rules (Russell, 1948; Popper, 1959; Bartley, 1962; Kuhn, 1962; Feyerabend, 1975); though pragmatic consequences resulting from effective action reinforce the search for better rules.

It should be clear that the process of seeking objective explanations and empirical rules is the essence of *erklären*.⁵ In its extreme pure form, *erklären* therefore can be seen as an impersonal, mechanical type of knowledge, lacking in feeling and personal involvement.

However, it should be noted that the pure form of *erklären* as 100% explicit knowledge is seldom, if ever, encountered. A major portion of a scientist's professional socialization is devoted to acquiring tacit knowledge and know-how so the future scientist can read between the lines and know things that are not actually stated in texts and articles (Heimdahl, 1974). The student's first experiments in the laboratory are highly controlled by rules and instructions. This is cookbook science, very mechanical and without much of a feeling for how science really works. As the student becomes enmeshed in experiments and studies that confront him with real life feedback, he begins to be shaped beyond the rules. Because rules are inherently inadequate, beginning students frequently have problems and make mistakes. However, the rules

have mentioned. Contingency-shaped personal knowledge gives one a rich, intuitive sense or "feeling" for when a given rule is inadequate, incorrect or should be broken. The rule inventor is responding from personal knowledge (not from a repertoire of "internalized rules") when he is creating or bending rules for others (or for the sociologist who may be studying him).

⁵ *Erklären* usually implies objective, explicit explanation. It is not restricted to logical positivism or operationism.

do orient students in the right general direction and make it possible for the contingencies of the scientific practica and the teacher's feedback to condition tacit knowledge. The mechanicalness of the early rule-governed behavior is smoothed out and shaped toward greater sensitivity as the student gains the tacit knowledge that no rules can communicate. As the student moves deeper into his professional socialization and into a career of research, he leaves the area where clear-cut rules have been formulated already and comes increasingly under the influence of new contingencies that may shape him in directions not explored by others, and shape tacit knowledge that is novel and heretofore unknown. Increased personal experience brings hunches and feelings that may eventually lead the scientist to a breakthrough that can then be encoded as explicit empirical knowledge—the new discovery.

Each generation of scientists steps into the causal stream at a different point, learns different tacit knowledge and rewrites the explicit knowledge called "science" to fit its current needs. Even within a given generation, there are many divergent schools or paradigms that expose their students to different shared examples, descriptive languages and criteria of adequacy, and hence condition different tacit and explicit knowledge (Heimdahl, 1974). Communication across paradigms can be difficult if members of different schools have been conditioned so differently that their verbal behavior does not strike a responsive chord in listeners from other paradigms.

It is not likely that scientists would contribute much to improving empirical rules if they operated exclusively within the domain of explicit knowledge, literal definitions and fully operational variables. "... [T]he ideal of eliminating all personal elements of knowledge would, in effect, aim at the destruction of all knowledge" (Polanyi, 1966:22). Reiterating the explicit knowledge of the previous generation is a starting point for learning science and exploring new questions, but it may not lead much beyond last-generation knowledge. The rule user must generate scientific behavior—in the form of hypotheses,

experiments, observations or assertions—which exposes both his tacit and explicit knowledge to natural (including social⁶) contingencies in order that his behavior and knowledge can be shaped beyond the point where the last-generation rules left off.

SCIENCE AND TECHNOLOGY

In some ways, sociologists operate under the same constraints as scientists in other disciplines. Last-generation knowledge is transmitted by rules; and new knowledge arises in large part from (1) tacit member's knowledge, shaped by long experience with fellow humans, social networks and institutions, and (2) tacit scholarly knowledge shaped by a professional socialization with participant observation, experimental findings, survey data or historical analysis. As in other sciences, the tacit aspect of new knowledge must be encoded in verbal (or mathematical) formulations before it can be shared with others. The resulting public knowledge falls short of capturing the richness and sensitivity of the original personal knowledge; but the advantages of communication, rapid education, prediction and control compensate to a degree.

In addition, there are special constraints on sociologists. Since sociology focuses on the activities of human beings, who have feelings and personal knowledge themselves, explicit knowledge seems especially inadequate for capturing the richness of social phenomena. Thus, it is understandable that sociologists have emphasized the value of personal experience for transcending the limitations of explicit sociological knowledge. Rather than relying solely on less personal knowledge, students of social behavior benefit by gaining a feeling for what it is to be shaped by the contingencies that affect their subjects.

⁶ Most scientists have their scientific knowledge shaped by dual contingencies: feedback from experiments or observations on the target of their research, and feedback from colleagues about their ideas (verbal behavior). The idealized view of science portrays science as relying primarily on the former; but social negotiation, learning how to sell an idea, debunking other positions, and other social moves are all part of doing science.

In the remainder of this section we will consider the similarities between sociology and the other sciences; and in the next sections we will focus on special issues that arise when studying humans.

If a sociologist has lived through the contingencies of being in a penitentiary, he knows what it feels like better than if he had read a dozen books on the topic. But when it comes time for the sociologist to pass on his wisdom to others who have not lived through the real contingencies, he faces a problem. (1) If he wants to help others gain as much personal knowledge as possible, he would tell his students to live through a prison sentence themselves. Let the contingencies shape their behavior and feelings as only a true penitentiary experience can. But this strategy is time consuming, and the students would still never arrive at the teacher's unique knowledge since the students' experiences doubtlessly would differ in many ways. (2) If the sociologist gave the students a ten-day free trial visit in the pen, the students might arrive at a modicum of personal experience,⁷ but this method falls short of producing the effects of a several-year conditioning experience. (3) Writing another book about what it is like in the pen would probably fall short of the ten-day free trial for communicating personal knowledge; but its attractiveness is obvious.

Personal knowledge resists easy communication: people have to share similar contingencies in order to learn similar feelings and knowledge. Two people who have both lived through similar experiences can talk about their feelings with great rapport; but the shared feelings that arise while discussing twenty days in solitary come from their both having lived through similar conditioning and not from one person's tacit knowledge being passed from his head to the head of the other person via a verbal medium. Words have little or different impact on the person who has not lived through similar experi-

ences.⁸ It is only when two people have a body of similar experiences that their words may strike a "responsive chord" (Schwartz, 1973).

That which we put into the communication has no meaning in itself. The meaning of our communication is what a listener or viewer *gets out* of his experience with the communicator's stimuli. The listener's or viewer's brain is an indispensable component of the total communication system. His life experiences, as well as his expectations of the stimuli he is receiving, interact with the communicator's output in determining the meaning of the communication. (Schwartz, 1973:25)

Because personal knowledge is not easily communicable (except to someone else who already has the same or similar tacit knowledge due to similar experiences), the sociologist finds himself in the same situation that reinforces scientists in other disciplines for creating explicit knowledge. The easiest means of transferring personal knowledge into a form that other people can learn from is to create symbolic descriptions and generalizations from prior experiences. This is explicit knowledge which can be used as rules by other people. It is easy for the sociologist's description of a person's behavior to sound like a set of rules that an actor could use to negotiate the scenes and events of the person's world. The rule quality of the sociologist's description does not imply, however, that the people he studied actually use these rules. Their knowledge is frequently tacit and quite different from the sociologist's explicit description (see fn. 5).

As do scientists in other fields, sociologists find that their rules become more effective and can be used more sensitively by educating students and colleagues in the unspoken knowledge that allows them to read between the lines and understand more than was written. The person who reads the sociologist's findings does not gain all the feeling and personal knowl-

⁷ Zimbardo (1972) showed that six days in a simulated jail environment can produce many of the effects of real incarceration; but it is doubtful if a seasoned inmate would consider the simulation as an adequate substitute for his own experience.

⁸ Staats (1968) presents several experiments that show how a given set of words can elicit different emotional responses in different subjects, depending on their past Pavlovian conditioning with those or similar words.

edge that the writer has access to; but he benefits from the distilled work. If the reader becomes involved with experiences or research similar to those of the sociologist whom he read, the extra contingencies will begin to give him increased tacit knowledge for the issues.

THREE TYPES OF KNOWLEDGE

In order to show how sociology differs from the "hard" sciences, we will discuss the reasons why knowledge about human social behavior and action differ from other kinds of knowledge. The domain of human knowledge can be divided into an arbitrary number of subsets. For the purposes of this analysis, we will categorize human knowledge into three nested sets: knowledge about (1) any conceivable phenomenon, (2) human behavior and (3) action and attached meanings. Each of the latter sets is a subset of the prior set. Both tacit and explicit knowledge can occur in any of the three sets of knowledge.

The first two sets traditionally are associated with *erklärte* knowledge and the third set has been generally assumed to be reserved for *verstehen* alone. The present analysis, if correct, will alter previous conceptualizations of *verstehen* and *erklären*. *Verstehen* will be seen as the application of tacit knowledge in the third area, the set of subjectively meaningful and intended action. Thus, *verstehen* shares many features in common with other forms of tacit knowledge. Modern behaviorism allows *erklären* to be extended from its traditional domains (sets 1 and 2) to include explicit analyses of the third domain, meaningful and intended action.

1. *Knowledge about any conceivable phenomenon.* This is the most general and all inclusive type of knowledge. Polanyi and Skinner usually operate at this level of analysis. Any experience with physical objects, living organisms or human beings produces tacit knowledge to the degree that the experience is contingency-shaped; and it produces explicit knowledge to the degree that the experience is rule-governed. Mixed knowledge, lying between the extreme types, is probably more common than either extreme.

2. *Knowledge about human behavior.*

Even though we cannot know what it feels like to be an inanimate object, plant or nonhuman animal, we have access to inside information on how it feels to be human. We all know what it is like to be hurt, to be angry, to be happy. But how does a dog feel when it smells the urine of a strange dog that has marked in its home range, or a gorilla feel when the troop's silverback male gives his tattoo? To some degree, our being human gives us extra information about other humans that we cannot have for other species or for objects. Let us call this level of extra knowledge "empathy" during the remainder of the discussion. Empathy is usually conceived of as a type of tacit knowledge, a felt, subjective knowledge; but explicit empathy occurs when one person "takes the words right out of another's mouth" or comes up with identical verbal formulations of certain events.

Empathetic information is never perfect, since no two humans have identical genetic, nutritional and learning backgrounds, and hence are biologically and psychologically unique.⁹ But that does not negate the fact that being human gives us inside information about some of the private experiences that other humans have. Empathy can be placed on a continuum: the more background elements two people share in common, the more likely it is that they will be able to empathize with each other's emotional and cognitive experiences. People sometimes think that they can have complete empathy with others, but sharing identical feelings is an extreme of the continuum that is rarely approached.¹⁰

Both laymen and sociologists benefit from the extra information that is available from empathy. When a mother says, "How would you feel if someone stole

⁹ Biological differences introduce barriers to empathy. It is difficult for the average person to empathize with the experience of a blind person, a hunchback or a paraplegic.

¹⁰ Our society often generously reinforces people who say they are empathetic; and it is possible for a person to learn the verbal behavior of saying and thinking that he has much more empathy than is in fact the case. Because the verbal community cannot check the accuracy of one's thoughts or feelings, a person may be very incorrect in his understanding of others and yet label himself as empathetic.

your toys?" she is preparing the child to use empathy, and appreciate that other people might feel sad. Learning this type of empathy can be an important ingredient in developing self-control and moral behavior (Bandura, 1971). Through his writing the poet communicates the joy he feels while watching nature. The communication need not be perfectly accurate or photographic to be effective. The poet describes his sunset and we see ours; but as fellow human beings, we sense we are sharing something in common. Students of human behavior also benefit from empathetic knowledge. During experiments, participant observation or questionnaire analysis, the scientist often relies on his member's knowledge to fill in gaps, read between the lines, suggest hypotheses and direct his research effort. The member's empathetic knowledge is never perfectly accurate, but it can be reasonably good if the researcher and his subjects have had many shared contingencies resulting in similar thoughts and feelings. Since science usually benefits from the attempt to create explicit information which is available for public verification, the scientific method points to the advisability of evaluating empathetic knowledge in as many ways as possible in order to have some measure of its accuracy.

Verstehen

3. *Knowledge about action.* Many sociologists follow Weber's orientation and focus special attention on the subset of human behavior called action, in which the individual attaches subjective meaning to his behavior. Especially important is social action, in which intended meaning is attached to socially directed behavior (Weber, 1925).

Since people do not attach meaning to all their behavior (especially reflexes and overlearned habits) or intend meaning in all their social interactions, action is a subset of the totality of human behavior. Because attaching meaning and intending are often private behavior, it can be exceptionally difficult to judge when these activities are taking place. Two people can attach different meanings to the same act, and a third person attach no meaning to it.

Observers can argue for long periods about whether a person actually intended to commit a crime or hurt another person's feelings. It is especially difficult to understand the meanings and intentions of people from other cultures or from significantly different subcultures.

The Weberian method for understanding action is through *verstehen*. Since Weberian action is only a subset of human behavior, *verstehen* is seen here as a subset of empathy. Thus, *verstehen* shares characteristics in common with empathy (as described above), but it has additional unique features. Abel (1948) argued that our understanding of other people's action is based on generalizations from direct personal experience and assumes that the emotions of others function similarly to our own. More recently, Munch (1957) and Wax (1967) added that the route to *verstehen* is shared experience. The more closely two people share the same socialization histories, the more they will have an accurate understanding of each other's feelings, meanings and intentions. The only way to approach this personal type of understanding for people in other cultures or subcultures is to go live with them: the more complete one's resocialization in a new group, the greater the possibility for *verstehen*.

This position basically is compatible with the behavioral position on *verstehen*. Two people who have shared similar contingencies of reinforcement will be conditioned to have similar feelings, thoughts, intentions, meanings and so forth.¹¹ Nat-

¹¹ Munch (1957) stresses that "*my own behavior . . . my own experience . . . is ultimately the only source from which I can have empirical knowledge . . . of mental conditions and events in other persons.*" The behaviorist would agree completely. Munch goes on, however, to say that a person infers by analogy that other people would feel as he does if they are involved in analogous activities, though the inferences are not necessarily made at a conscious level of thought. Behaviorists do not deny that people can infer by analogy, but they would point out that most cases of empathetic feeling and *verstehen* do not depend on such highly cognitive procedures as are indicated by inference by analogy. For example, at a funeral, the family cries for the lost member. They cry and grapple with the meaning of death as a response to the total stimulus configuration. Friends of the family are also exposed to many of the same stimuli, and the stimuli elicit similar thoughts, mean-

urally, no two people have totally identical socializations; hence perfect verstehen is impossible. This does not, however, negate the value of attempting to gain as much empathetic understanding and verstehen as possible. It merely points out the limitations of the method and provides objective criteria for establishing the areas in which subjectively felt empathy and verstehen may not be valid sources of information.

The sociologist who wishes to capitalize on Weberian verstehen is most likely to arrive at understanding which is valid if he restricts his studies to his own subculture or becomes resocialized into a new subculture by doing participant observation (Wax, 1967). The more the sociologist's experiences approach the learning experiences of the people he studies, the more likely his understanding of their actions is to be valid, though perfect validity can never be attained. As Martin (1968) and Scott (1971) point out, the intellectual socialization of the sociologist tends to separate him from many of those he would like to study; and this frequently leads sociologists to err by reading into other people's behavior a variety of cognitive, moral and intellectual processes that may not in fact exist. Thus, the ex-con later trained as a sociologist might do a better job at a verstehende sociology of prisons than would a college-educated scholar with no penitentiary experience.¹² To the degree that one wishes to benefit from verstehen, it is imperative to live it, to get one's fingers dirty and to become enmeshed in the contingencies that affect those whom one studies. As a consequence verstehen in historical analysis is

ings and emotional responses, especially for those who have recently shared the experience of losing a loved one, too. The more similar their past experiences with death have been, the greater their empathy and verstehen will be. In contrast, a child might cry merely because so many other people were crying, and not because he understood the meaning of sorrow or death.

In empathetic understanding and verstehen, each person's response is a direct consequence of his own past conditioning experiences. Inferring by analogy is superfluous, and rings more of explicit, erklärte knowledge than of verstehen.

¹² As an example of a sensitive, autobiographical verstehen of prison and crime that few professional sociologists could have written, see Braly (1976).

most accurate if one studies people and institutions within one's own culture, value systems and paradigms. "The more radically they differ from our own ultimate values, however, the more difficult it is for us to make them understandable (verständlich) . . ."; ". . . though sometimes we are able to grasp them intellectually" (Weber, 1925:91).

There is a large behavioral literature analyzing the conditions under which people learn to attach meaning to objects, persons, symbols, thoughts and actions (e.g., Skinner, 1945; 1957; 1969; 1974; Bem, 1964; 1970; Staats, 1968; Mahoney, 1974). This literature reflects the concern of modern behaviorists for extending a scientific (erklärte) understanding of human behavior into the realm of subjective events.¹³ The behavioral analysis concurs with Wittgenstein's (1953) position that "the meaning of a word is its use in the language."¹⁴ In addition, behaviorists seek to trace word usage, along with the behavioral and emotional elements related to words, back to the social and nonsocial environments from which these responses were learned. The social environment provides words, rules, prompts, models of usage, and contingent feedback which aid the individual in learning the discriminations and meanings used in his culture. Instead of merely assuming that the norms, meanings, attitudes and significant symbols of a culture are simply internalized, the behaviorist attempts to specify the learning contingencies involved in the complex processes of internalization.¹⁵

¹³ Dilthey (1927), Weber (1925), Abel (1948), Munch (1957), Wax (1967) and others have all tried to give explicit (erklärte) descriptions of how verstehen takes place. The attempt at an erklären of verstehen is not new.

¹⁴ "Wittgenstein and Skinner are very much alike in their analysis of the nature of meaning itself. For both, there are no such *things* as meanings, where meanings are taken to be mental entities somehow focally involved in communication. For both, a search for meaning can lead only to the study of word usage, to the analysis of verbal behavior as it is actually seen to take place. For both, the meaning is the usage." (Day, 1970:368). Day describes ten fundamental areas in which Wittgenstein and Skinner concur in their analysis of language.

¹⁵ Schrag (1961) and Scott (1971) point out that the

During each individual's history of conditioning—much of which is social conditioning—the person learns to attach meaning to stimuli (such as objects, people, symbols and actions). The sight of a bee or the sound of the word “love” has little or no meaning to a very young child who has not had learning experience with these stimuli. Through operant and Pavlovian conditioning, the child will learn from models, prompts, rules, and contingent feedback to attach verbal, nonverbal and emotional responses to the stimuli. Observers may infer that a person knows the meaning associated with a stimulus if an appropriate response from any of the three categories—verbal, nonverbal or emotional—is shown. For example, if a child displays strong emotional responses when using the word “love” and behaves in a gentle, affectionate manner with those he loves, the observer would probably conclude that the child attaches meaning to the word, even if the child could not give an articulate verbal explanation of the meaning of love. A deeper or richer level of meaning is often inferred, however, if a person has learned a harmonious melange of all three types of response—verbal, nonverbal and emotional.

On the other hand, if any or all the responses appear incongruous or inappropriate to the stimulus (as judged from the cultural perspective of the observer), the person might be judged as immature, “mentally disturbed” or at least unaware of the appropriate meaning of the stimulus. A man who verbally claims to attach a great deal of meaning to love and affection—but whose cruel, selfish be-

havior indicates otherwise—might not be judged to know the true meaning of love.

The behavioral analysis of meaning focuses on the antecedent history of conditioning that has given a stimulus complex the ability to evoke verbal, nonverbal and emotional responses. In order to analyze the insensitive man's meaning of love, it would be necessary to trace the conditioning of all three aspects of his behavior. Because the contingencies and rules that produce verbal, nonverbal and emotional responses can be relatively independent, it is possible for people to learn incongruous, even hypocritical response patterns (Scott, 1971). When the contingencies reward saying one thing and doing another—professing enduring love, then cashing in on one-night stands—discrepancies can appear between the verbal and nonverbal meanings a person attaches to the concept “love.” In addition, the emotions attached to what the hypocritical lover says and what he does can also be quite different. The behavioral analysis is thoroughly sociological because it traces the causes for attached meaning back to the social environment in which the person learned to attach and use meaning. Social learning theory, which has produced extensive empirical data on the effects of models, prompts, rules and contingent feedback (cited in Bandura, 1969; Mahoney, 1974), provides an ideal tool for sociologists who seek to link social behavior to societal causes.

Naturally, a person may attach different meanings to a given stimulus in different contexts or at different ages in life. A person may learn to discriminate that promises of love do not all carry the same meaning, in the sense that not all are associated with the same kind of relationships. As a person gains greater discriminative ability from firsthand experience or rule-guidance, the person will attach different meanings to the same verbal stimuli, and rely on context cues, nonverbal indicators or past experiences to make the discrimination. The way the person responds to these context cues will depend on the prior contingencies of reinforcement. “. . . [T]he contingencies, not the mind, make discriminations” (Skinner, 1974: 105). Thinking “does not explain overt be-

concept “internalization” is extremely vague and tells us little. If someone said that a Martian internalized a hamburger, we would not know anything more than that the hamburger went from the outside to the inside. We would not know whether Martians have proboscises, mouths, little trap doors for inserting food through the side or whether food entered the opposite end from ours. This lack of knowledge makes it difficult to understand, predict or control most of the internalization process. Behaviorism allows us to specify how cultural practices are learned from the social environment and hence to make an *erklärte* science of the internalization and use of symbols, meanings and values in human action.

havior: it is simply more behavior to be explained" (Skinner, 1974:104). Tracing the discrimination back to the person's past history of social learning is basically a sociological enterprise, locating the causes of social behavior in the social environment.

It is possible for people to attach meaning without giving any outward behavioral manifestations of the event. This type of private behavior poses a problem for both *verstehende* and *erklärte* knowledge of other people's actions. The problem of intersubjectivity is insuperable. Yet all of us can observe our *own* private behavior and thus test behavioral principles with a limited sample. In addition, we can study in others (i) the contingencies that control the silencing of overt verbal behavior and (ii) predicted changes in overt behavior that result from experimentally induced changes in private behavior. The controlling variables for private verbal behavior lie in the environment; and they are subject to empirical study and control. For examples of empirical research that demonstrate techniques for studying the environmental factors that control private verbalizations and for measuring the predicted changes in subsequent overt behavior, see Homme (1965), Bem (1967), Blackwood (1970) and Meichenbaum and Cameron (1973; 1974). The data from these types of studies indicate that private verbal behavior is no different from overt verbal behavior, except that one set of contingencies and rules makes it silent. These data provide empirical support for the behavioral contention that there is no reason to make a mystery out of thinking, merely because it eludes direct empirical measure.

The goal of modern behaviorism is to specify the contingencies by which subjective processes are "internalized." The behavioral analysis traces the act of attaching meaning back to the environment, especially to socializing experiences, rather than leaving the act of attaching meaning up to the individual's "free will." Many sociologists flatter their readers by giving their actors elbow room to define or redefine the world, as if the constraints of biology, learning and social realities were not 100% binding. The behavioral analysis

stresses that free will cannot be studied,¹⁶ that the socialization of attaching meaning can be studied, and that the role of sociologists is to determine how our social environments condition us to act as we do.¹⁷ Private behavior is not off limits to the modern behaviorist (Homme, 1965), and solid empirical work (summarized in Mahoney, 1974) already has been done to illuminate the role of internal activities in the causal flow of behavior.

The scientific analysis of personal knowledge, empathy and *verstehen* has important humanistic implications. Spiritual leaders, demagogues and other gurus often mobilize social movements using the authority of their own feelings and/or calling upon our feelings, empathy and *verstehen* to guide us into following them. When the followers of Hitler, Reverend Moon and others apply the test of subjective feeling, their commitment feels right. However, *any* commitment will feel right or good if it has been paired with strong reinforcement; hence the test of feeling tells only what one person found rewarding, not necessarily what will best benefit the whole society.

Social action based upon a scientific analysis of human behavior is much more likely to be humane. It can be transmitted from person to person and epoch to epoch, it can be freed of personal predilections and prejudices, it can be constantly tested against the facts, and it can steadily increase the competence with which we solve human problems. (Skinner, 1969:167)

CONCLUSION

Tacit and explicit knowledge represent opposite poles of a continuum. During everyday life and during scientific investi-

¹⁶ Because it would be an uncaused cause, free will (if it existed) would follow no laws and hence never be amenable to empirical study. Even humanistic sociologists have noted that free will creates "*a priori* difficulties within the framework of a sociological argument . . . [because] freedom is not empirically available" (Berger, 1963:122).

¹⁷ Scott (1971) argues that the role of the sociologist is to trace the internalization of norms, moral values and other mentalistic states back to the natural social contingencies that produce them. He provides an excellent example of how this type of analysis can be done in the area of moral behavior.

gations, people rely on complex mixtures of methods all along the tacit-to-explicit knowledge continuum, melding hunches and feelings at one point with instructions and rules at another. Analytically, it is easier to describe the poles of the continuum (on any of the three levels of knowledge discussed earlier) than all the complex gradations in between; but the intermediate methods are more common than the extreme cases. When solving a problem or conducting research, most people combine a large repertoire of tacit knowledge with rules, procedural guidelines and time-tested principles.

That which is true for tacit and explicit knowledge as a generality is true for the third level described above, *verstehende* and *erklärte* knowledge of action and social action. Debates about *verstehen* and *erklären* often polarize the two methods rather than presenting them as differently weighted mixtures of tacit and explicit knowledge. Science is often described as a completely rule-governed enterprise, as if scientists followed the rules of induction, deduction, axiomatic proofs and analytic logic, without recourse to hunches and intuitions. And those who rebel against science often claim the value of natural knowledge, which supposedly does not rely on rules, equations, or scientific laws. The polar extremes stand out as black and white, and people often succumb to making black and white arguments rather than entering the more complex territory that lies in between (Kaufmann, 1973).

It is time for a synthesis. Both tacit and explicit knowledge are natural parts of human behavior. When it comes to understanding the meaning that people attach to their actions, all combinations of *verstehen* and *erklären* can produce useful knowledge. Some of us may lean toward one end of the continuum or the other, but it is unlikely that anyone ever approaches a complete dependency on either extreme. Practicing scientists acknowledge the value of tacit knowledge in their pursuit of empirical laws: "There is no logical path to these laws; only intuition, resting on sympathetic understanding of experience, can reach them" (Einstein, 1918:226). And societies without science are not devoid of rules in the form of folklore, tribal

customs and traditions. It is time to end unnecessary hostilities between the two modes of knowing, even in the third area of understanding subjective and meaningful events.

REFERENCES

- Abel, T.
1948 "The operation called *verstehen*." *American Journal of Sociology* 54:211-8.
- Bandura, A.
1969 *Principles of Behavior Modification*. New York: Holt.
1971 "Vicarious and self-reinforcement processes." Pp. 228-78 in R. Glaser (ed.), *The Nature of Reinforcement*. New York: Academic Press.
- Bartley, W. W.
1962 *The Retreat to Commitment*. New York: Knopf.
- Bem, D. J.
1964 *An Experimental Analysis of Beliefs and Attitudes*. Ph.D. dissertation, Department of Sociology, University of Michigan.
1970 *Beliefs, Attitudes, and Human Affairs*. Monterey: Brooks/Cole.
- Bem, S. L.
1967 "Verbal self-control: the establishment of effective self-instruction." *Journal of Experimental Psychology* 74:485-91.
- Berger, P. L.
1963 *Invitation to Sociology: A Humanistic Perspective*. Garden City: Doubleday.
- Blackwood, R. O.
1970 "The operant conditioning of verbally mediated self-control in the classroom." *Journal of School Psychology* 8:251-8.
- Braly, M.
1976 *False starts: A Memoir of San Quentin and Other Prisons*. Boston: Little, Brown.
- Bridgman, P. W.
1952 *The Nature of Some of Our Physical Concepts*. New York: Philosophical Library.
1959 *The Way Things Are*. Cambridge, Ma.: Harvard University Press.
- Commoner, B.
1971 *The Closing Circle*. New York: Knopf.
- Day, W. F.
1970 "On certain similarities between the philosophical investigations of Ludwig Wittgenstein and the operationism of B. F. Skinner." Pp. 359-76 in P. B. Dews (ed.), *Festschrift for B. F. Skinner*. New York: Appleton-Century-Crofts.
- Dilthey, W.
1927 *Wilhelm Diltheys Gesammelte Schriften, VII. Band: Der Aufbau der Geschichtlichen Welt in den Geisteswissenschaften*. Leipzig: Verlag von B. G. Teubner.
- Einstein, A.
[1918] "Principles of research." Pp. 224-7 in A. Einstein, *Ideas and Opinions*. New York: Crown.
1954

- Feyerabend, P.
1975 *Against Method: Outline of an Anarchistic Theory of Knowledge*. London: Humanities Press.
- Garfinkel, H.
1967 *Studies in Ethnomethodology*. Englewood Cliffs: Prentice-Hall.
- Heimdahl, C.
1974 "The structure of scientific revolutions: Kuhn and his critics." *Communication and Cognition* 7:33-60.
- Homme, L. E.
1965 "Perspectives in psychology: XXIV. Control of coverants, the operants of the mind." *Psychological Record* 15:501-11.
- Huxley, A.
1954 *The Doors of Perception*. New York: Harper and Row.
- Kaufmann, W.
1973 *Without Guilt and Justice*. New York: Wyden.
- Kuhn, T. S.
1962 *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Mahoney, M. J.
1974 *Cognition and Behavior Modification*. Cambridge, Ma.: Ballinger.
- Martin, M.
1968 "Understanding and participant observation in cultural and social anthropology." Pp. 303-30 in R. S. Cohen and M. W. Wartofsky (eds.), *Boston Studies in the Philosophy of Science*, Vol. 4. Dordrecht, Holland: Reidel.
- Meichenbaum, D. and R. Cameron
1973 "Training schizophrenics to talk to themselves: a means of developing attentional controls." *Behavioral Therapy* 4:515-34.
1974 "The clinical potential of modifying what clients say to themselves." Pp. 263-90 in M. J. Mahoney and C. E. Thoresen (eds.), *Self-Control: Power to the Person*. Monterey: Brooks/Cole.
- Munch, P. A.
1957 "Empirical science and Max Weber's *verstehende soziologie*." *American Sociological Review* 22:26-32.
- Polanyi, M.
1959 *The Study of Man*. London: Routledge.
1960 *Personal Knowledge*. Chicago: University of Chicago Press.
1966 *The Tacit Dimension*. Garden City: Doubleday.
- Popper, K.
1959 *The Logic of Scientific Discovery*. New York: Basic Books.
- Reynolds, G. S.
1968 *A Primer of Operant Conditioning*. Glenview: Scott, Foresman.
- Russell, B.
1948 *Human Knowledge: Its Scope and Limits*. New York: Simon and Schuster.
- Schrag, C.
1961 "Some foundations for a theory of correction." Pp. 309-57 in Donald Cressey (ed.), *The Prison: Studies in Institutional Organization and Change*. New York: Holt.
- Schwartz, T.
1973 *The Responsive Chord*. Garden City: Doubleday.
- Scott, J. F.
1971 *Internalization of Norms: A Sociological Theory of Moral Commitment*. Englewood Cliffs: Prentice-Hall.
- Skinner, B. F.
1945 "The operational analysis of psychological terms." *Psychological Review* 52:270-7, 291-4.
1953 *Science and Human Behavior*. New York: Macmillan.
1957 *Verbal Behavior*. New York: Appleton-Century-Crofts.
1969 *Contingencies of Reinforcement*. New York: Appleton-Century-Crofts.
1974 *About Behaviorism*. New York: Knopf.
- Staats, A. W.
1968 *Learning, Language, and Cognition*. New York: Holt.
- Thoresen, C. E., and M. J. Mahoney
1974 *Behavioral Self-Control*. New York: Holt.
- Wax, M. L.
1967 "On misunderstanding *verstehen*: a reply to Abel." *Sociology and Social Research* 51:323-33.
- Weber, M.
[1925] *The Theory of Social and Economic Organization*. Ed. by T. Parsons. New York: Oxford University Press.
1947
- Wittgenstein, L.
1953 *Philosophical Investigations*. New York: Macmillan.
- Zimbardo, P. G.
1972 "Pathology of imprisonment." *Transaction/Society* 9(6):4-8.

EXPLAINING THE LOW USE OF HEALTH SERVICES BY THE POOR: COSTS, ATTITUDES, OR DELIVERY SYSTEMS?*

DIANA B. DUTTON

Stanford University

American Sociological Review 1978, Vol. 43 (June):348-368

The poor, especially children, continue to receive fewer health services relative to need than the affluent. Explanations have traditionally focused on cost constraints or on cultural differences. This paper provides empirical evidence that such explanations do not fully account for income trends in preventive and symptomatic use. A third explanation, based on inadequacies in delivery systems used by the poor, is required. Factors representing each explanation are added sequentially to a multivariate model in order to shed light on their role in the income-use relationships. Particularly instructive are changes in estimates when types of delivery systems are added, since system differences have been ignored in much previous research. The importance of factors promoting use among the poor, e.g., public assistance, is underestimated, while the role of individual characteristics, e.g., attitudes, is overestimated. Results suggest that neither financial access nor health education, without accompanying improvements in delivery systems, will eliminate income differentials in use.

INTRODUCTION

Following Koos's report in 1954 that lower class persons were less inclined than the more affluent to consult a doctor for common illness symptoms, low rates of health care use by the poor have been widely recognized, and confirmed in national data as well as in many local studies. But over the years, income gaps in utilization have steadily narrowed. In 1970, according to national statistics, the traditional direct relationship between income and use was replaced with a U-shaped pattern (National Center for Health Statistics [NCHS], 1974b:5). Since then, the lowest physician visit rates have been in the middle-income range—indeed, rates among the poor are now *higher* than those of the nonpoor (5.6 vs. 4.9 visits per person per year, respectively, in 1973, for

example; Wilson and White, 1977). Does this mean that the goal of equal access has finally been achieved?¹ Can low use by the poor be relegated to an unfortunate chapter in America's social history? The answer to these questions, regrettably, must be no, despite claims to the contrary (Roghamann, 1974; Monteiro, 1973). Many income-related differentials still exist, when appropriate measures are considered.

There has been much disagreement about the cause of these use differentials (Kadushin, 1964; 1967; Goering and Coe, 1970; Salkever, 1975). Until only recently, explanations have focused predominantly on the financial constraints facing the poor

* I am grateful to Dr. David Kessner and the rest of the staff of the Institute of Medicine research group for making available and for assistance in interpreting the data sources on which the present study is based. A generous grant from the Milbank Memorial Fund supported most of the earlier phases of this research. I am indebted to many persons for help in developing the substance and approach of the research, most recently to Ralph Silber, who made a substantial contribution to this paper. An earlier version was presented at the annual meeting of the Pacific Sociological Association, Sacramento, 1977.

¹ Following Illich (1976), one might question whether increasing the access of the poor to professional medical care is a desirable goal. Yet, equity demands that at least the opportunity for care be equally distributed, with the number of services determined by health needs rather than ability to pay (Bryant, 1977). Moreover, there is much evidence—clinical comparisons (Kravits and Schneider, 1975), length of hospitalization stays (NCHS, 1966), and community surveys (Goering and Coe, 1970)—which suggests that many of the poor presently go without needed care. In short, doubts about the value of an unlimited quantity of medical care are appropriate, but should not lead to doubts about the need for increased opportunities for access to care among the poor.

or on their attitudes and priorities concerning health care. In this paper, I will argue that both of these explanations, even together, are inadequate. A third explanation is also necessary, one that implicates various barriers and structural impediments of the health care systems used by the poor. These three alternative explanations have very different policy implications: if costs are the problem, better health insurance is the remedy; if inappropriate health attitudes are the problem, health educational programs are the remedy; if inadequacies in the health delivery systems are the problem, structural improvements in these systems are the remedy.

Although there is considerable evidence for each of the three explanations considered separately, they have rarely been evaluated together. This paper describes the results of such an evaluation, for use of both preventive as well as symptomatic services. The findings provide empirical evidence in support of arguments that structural reform of existing health care systems is essential in order to improve access to care for the poor.

What Is the Relation between Income and Use?

The national statistics cited at the outset are misleading for at least three reasons (see Davis, 1976:127 for others). First, there is a strong positive relation between children's visits and income, which is hidden in statistics based on all age groups. High income children (\$15,000 + annual family income) averaged almost twice as many physician visits as low income children (below \$3,000) in 1971 (NCHS, 1974b:20), and disparities in the proportion without any visits in a given time period were even greater.

Second, statistics comparing physician visits by income level ignore the underlying distribution of medical need. There is considerable evidence that the poor experience more disability and lower levels of health than the nonpoor (NCHS, 1974a:6,25). Studies which have taken level of disability into account indicate that the poor still receive *less* care for their illnesses than the affluent, similar

visit rates notwithstanding (Davis and Reynolds, 1976; Aday, 1975). Thus, although access to care among the poor has clearly improved in recent years, one cannot infer that services are now distributed equitably relative to health needs.

Third, national data on number of physician visits confound visits of different types (merging preventive and elective care, for example) with clinically necessary therapeutic care. This is an important distinction, since these two categories of care, when examined separately, generally reveal opposite income trends. Measures of use which involve patient discretion—such as preventive care and the initial decision to seek symptomatic care—almost always display a strong *positive* relation to income (NCHS, 1974b:10; Bice, 1971; Kravits and Schneider, 1975). On the other hand, procedures identified as clinically mandatory and other indicators of therapeutic care usually show a strong *negative* relation to income (Andersen, 1975).² When elective and therapeutic services are combined, as they are in number of visits per person per year, these two opposite income trends are confounded. Unless the two types of care are kept separate, the question of what causes either trend can never be posed.

Three Different Explanations for Low Use Rates by the Poor

Before assessing the validity of alternative explanations offered to account for positive income trends based on data from the present study, let us first consider the arguments and evidence for each in more detail.

(1) The financial coverage explanation.

² Unlike the patient-controlled measures of discretionary-use, therapeutic services are controlled primarily by health care providers. Providers are clearly influenced to some extent in their prescribing practices by nonmedical considerations (Freidson, 1974; Mechanic, 1975; Dutton, 1977); but compared with patients' initial decisions to seek diagnosis, patterns of therapeutic care are more likely to depend on the clinical severity of the illness condition (Kovner et al., 1969; Richardson, 1970). Indeed, Zola (1966) has argued that patients' decisions to seek care may be virtually unrelated to the clinical severity of the symptoms.

According to this explanation, low use rates by the poor reflect cost constraints. The poor cannot afford to purchase the services they need—incomes are low, and adequate insurance against health care costs is lacking. This view was implicit in the Medicaid program, which addressed neither the predisposition of the poor to seek professional services nor the availability of such services. The assumption was that, given adequate dollars, the poor could express their demand for care in the medical marketplace, and that this demand would then attract a sufficient supply of health providers. In this "market model" approach to the health problems of the poor, financial access was expected to eliminate problems of physical access. As it turned out, this approach proved not only inadequate (U.S. Department of Health, Education, and Welfare, 1968), but also actually created some new problems of abuse by health providers (Goldberg, 1976). Nevertheless, Medicaid did result in higher use rates among the families covered, and, in this respect, did improve access (Aday, 1975).

(2) *The culture of poverty explanation.* The culture of poverty concept as formulated by Oscar Lewis in 1966 referred to a way of life that was both an adaptation and a reaction of the poor to their social and economic situation. This way of life produced a constellation of social and psychological traits which, handed down from generation to generation, served to perpetuate the poverty cycle. Inadequate use of health services, leading to poor health, was seen as one manifestation of this social maladaptation.

Such maladaptation might reflect alienation from society in general (Durkheim, 1933; Blauner, 1964) or from health care institutions in particular (Seeman and Evans, 1962; Suchman, 1964; Morris et al., 1966; Becker and Maiman, 1975). Other subcultural traits also have been linked to low use; these include a crisis-oriented approach to life (Rainwater, 1968), a greater willingness to put up with illness symptomatology (Koos, 1954), or simply a tendency not to define it as illness (Zola, 1966).

Despite much research on the role played by various attitudinal measures,

the evidence is inconclusive (McKinlay and Dutton, 1974). While some discrepancies in the reported findings may reflect methodological differences or deficiencies, doubt is also cast on the validity of the culture of poverty viewpoint itself—specifically, on its portrayal of beliefs and attitudes as predominant over economic and situational barriers to care. This viewpoint, summarized in the following quote from Rosenstock (1975:215–6), has direct policy implications:

The culture of poverty may originally have been based on a history of economic deprivation, but it seems to be a culture exhibiting its own rationale and structure, and reflecting a way of life that is transmitted to new generations. It is therefore suggested that while *financial costs may serve as barriers to obtaining health services, their removal would probably not have the effect of creating widespread changes in health behavior of the poor, at least not in the foreseeable future.* (emphasis added)

The culture of poverty concept has been, of course, widely attacked (Riessman, 1974; Leacock, 1971; Valentine, 1968). Yet many investigators continue to stress the importance of cultural factors as determinants of utilization (Becker et al., 1977b). Berkanovic and Reeder (1974), for example, have urged that more attention be given to the "differential ordering of problems and priorities" and the "diverse value systems" of the poor and less to financial barriers they might face. The implication is that these cultural factors, rather than financial or structural access barriers, are at the root of low use rates among the poor.

(3) *The system barriers explanation.* The third explanation focuses on various inadequacies of the health care systems typically used by the poor. To begin with, access itself is often difficult. Physicians tend to be scarce in poverty areas (Bullough, 1972; May, 1975), and travel difficulties are often compounded by inadequate transportation (Weiss and Greenlick, 1970). Access problems frequently continue in the form of long waits in the provider's office or clinic (Aday, 1975; Bice, 1971).

But beyond access lies a more fundamental problem: a dual system of medical

care, in which the poor utilize "public" sources—hospital outpatient departments, emergency rooms, and public clinics—while middle and upper income groups utilize private sources—physicians in solo or group practice (NCHS, 1972). In the so-called "public" sources (an ironic name in light of their frequently high charges), organizational problems are commonplace. Patients must often maneuver between multiple clinics to obtain basic primary care services, and these services are generally disease-oriented rather than preventive. Further, the atmosphere in such institutions is often dehumanizing. To the low income patient, the institutions may seem, in Strauss's words (1969:152):

... terribly massive and complex, crowded and busy; while the personnel seem often impersonal, brusque, or even insulting[;] ... the physicians go from patient to patient, spending brief moments with most [;] ... patients may sit for long periods of time waiting to be called. ... Patients see all of this, and may simply respond fatalistically to the rush and bustle.

Notice that Strauss suggests that the responsibility for the alienation of patients lies at least as much with the "culture of medicine" (Levine, 1969) which dominates these institutions as with the cultural background of the patients. Thus, the system barriers explanation assigns primary responsibility for low use rates to various inadequacies of the health care *systems* used by the poor. Low utilization is viewed as the natural response to multiple negative experiences with these systems (Riessman, 1974).

each explanation are added to the analysis. The initial estimates may be either increased or decreased (or they may remain unchanged).³ If they are decreased, then the factors which have been added can be said to mediate (i.e., explain) the positive income-use relations initially estimated. If the initial estimates are instead increased when a group of factors is added, then these factors can be said to "confound" the income-use relationship apparent initially. Rather than explaining this initial relationship, they obscure it.

Using this approach, a multiple regression model of ambulatory care use is developed incrementally by adding successive groups of variables which represent each of the three explanations. The strength (size and significance) of the parameter estimates for income is compared after each group of variables is added, providing empirical evidence of the explanatory power of each explanation. In addition, the magnitude of the income differentials which remain at successive stages is calculated, based on the differences between the estimated use rates of the highest and lowest income groups. These income differentials, which are adjusted for all of the other variables in the model at each stage, provide another test of the success of alternative explanations in accounting for the income-use relation. Finally, some direct effects on use (standardized regression coefficients) that were estimated for each group of variables in the fully-specified model are also considered. These direct effects indicate the impact on use, independent of income, of selected variables in each set.

DATA AND METHODS

Evaluating Alternative Explanations for Positive Income Trends in Use

In each of the three explanations discussed, a different set of factors is portrayed as the critical link which explains the low rates of discretionary use by the poor. One way to evaluate these explanations is to observe how initial estimates of the income-use relation change as successive groups of factors representing

³ If they remain unchanged, then the factors added clearly do not mediate the income-use relationship. Whether the initial estimates for income will be increased or decreased depends on the relationship between income and the mediating factor (X) and also between X and use. If the following relation is estimated: (1) $Use = b_1 Inc + e$; while the true structure is: (2) $Use = b_1 Inc + b_2 X + e$ and (3) $X = b_3 Inc + e$; standard specification analysis algebra shows that $E(b_1^*) = b_1 + b_2 b_3$ (where E denotes the expected value). Notice that if $b_2 b_3 > 0$, $b_1^* > b_1$, and the initial estimates will be decreased when X is included in the analysis. If instead $b_2 b_3 < 0$, then the initial estimates will be increased when X is included.

Model Specification and Development

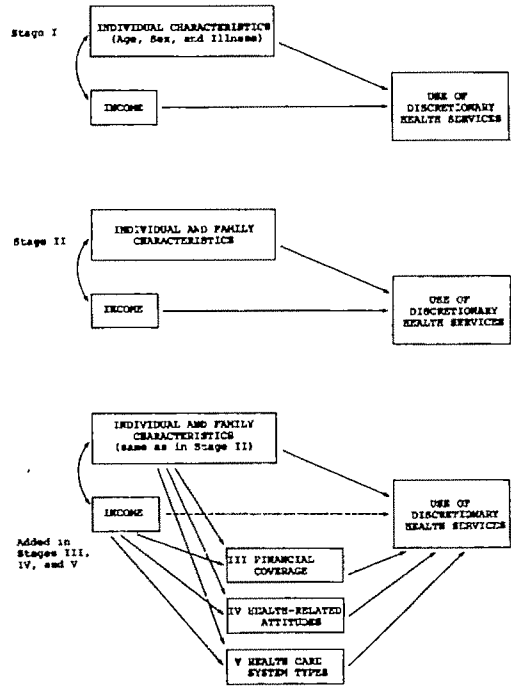
Two types of discretionary use are analyzed:⁴ preventive use, represented by children's and respondents' health check-ups; and *initiation of care*, represented by one or more visits to the usual provider in the last six months, children's (nonschool) vision and hearing tests,⁵ and consultation with a doctor about children's ear problems. The coding, number of cases, and mean values of each measure are listed in Appendix A.

The initial model of use, including only age, sex, and illness, provides estimates of income trends in use adjusted for differences in these basic demographic factors. The second stage adds a group of variables representing lifestyle indicators commonly associated with poverty—education, occupation, race, and family composition. These measures have been linked in previous research with variation in use (McKinlay, 1972; Andersen and Newman, 1973); thus, estimates from this stage indicate whether or not income trends in use exist *independent* of such lifestyle measures. The specific variables included in Stages I and II are listed in Appendix B, with their coding and mean values. (A detailed rationale and supporting evidence for selecting and specifying these variables, and the parameter estimates obtained for them, are provided in Dutton, 1976.)

The remaining stages then test the three alternative explanations for income trends in use. Stage III adds two variables representing financial coverage (public and private health insurance) as a test of the financial access explanation; Stage IV adds a set of five health-related attitudes as a test of the culture of poverty explanation; and Stage V adds variables indicating the type of health care delivery system used as a test of the system barriers explanation. The specific variables added are

⁴ Physician-controlled measures of use (follow-up care and medication) were also analyzed, but are not discussed in this paper (see Dutton, 1976, for these results). They generally display the expected negative relation to income.

⁵ Vision and hearing tests were correlated with eye and ear problems, respectively; this suggested that their purpose was primarily diagnostic or therapeutic rather than preventive.



Note: Curved arrows denote correlation and straight arrows denote hypothesized causal relations. The dotted arrow indicates the changing relationship between income and use at successive stages of model specification.

Figure 1. Development of the Model of Use of Health Services

listed in Appendix C and are discussed in more detail as the findings are reported.

Figure 1 illustrates the five stages of the model. Each of these stages represents an alternative causal model, since each includes successively more factors which might serve as intervening paths in the initial relationship between income and use.⁶ The dotted arrow between income and use represents the relationship (if any) between the two which remains at successive stages.

In effect, then, five separate equations were estimated for each use measure, each containing a new group of variables added to those already in the equation from preceding stages:

⁶ Of course, the income trends in all but the final Stage (Stage V) are misestimated strictly speaking, since earlier models do not include the full set of relevant variables. The direct effects of income, attitudes, and financial access measures are all discussed later, however, based on results from the full model.

$$\text{Use} = \sum_{i=1}^3 \beta_i X_i + \sum_{i=4}^{12} \beta_i X_i + \sum_{i=13}^{14} \beta_i X_i$$

STAGE I STAGE II STAGE III

$$+ \sum_{i=15}^{19} \beta_i X_i + \sum_{i=20}^{23} \beta_i X_i$$

STAGE IV STAGE V

where Use = one of the seven dependent variables listed in Appendix A.

Stage I: $\sum_{i=1}^3 \beta_i X_i$ = Illness and Demographic Variables, as listed in Appendix B.

Stage II: $\sum_{i=4}^{12} \beta_i X_i$ = Family Structure, and Socioeconomic Variables, as listed Appendix B.

Stage III: $\sum_{i=13}^{14} \beta_i X_i$ = Financial Coverage Variables, as listed in Appendix C.

Stage IV: $\sum_{i=15}^{19} \beta_i X_i$ = Health-Related Attitudes, as listed in Appendix C.

Stage V: $\sum_{i=20}^{23} \beta_i X_i$ = Usual Health Care System Types, as listed in Appendix C.

e=stochastic error.

The five system types were represented by a set of dummy (0-1) variables, with solo practice as the comparison base (omitted variable). Thus, system effects were estimated as the differences between use rates in each of the four included systems compared with those in solo practice.

Ordinary least squares (OLS) was used as the estimating method for each of the seven use measures.⁷ In general, stan-

dardized (beta) coefficients are reported rather than unstandardized regression coefficients in order to facilitate comparison across equations which have dependent variables with different ranges, and also, in a single equation, across independent variables with different ranges. Nonstandardized coefficients were used to calculate the estimated differences in use rates between the highest and lowest income groups, reported in Table 3. These differences were computed by multiplying the income coefficients estimated in successive stages times the difference between the highest and lowest income codes (=7).

Data Sources

The analyses are based on two data sets collected in Washington, D.C. in 1970-1971, under the auspices of the National Academy of Sciences (a household interview survey, and a survey of providers identified by the households as *usual sources of health care*). (These surveys are described in detail in Kessner et al., 1974.)

The household survey was conducted in two geographic areas of Washington, D.C., in order to obtain a broader range of income levels and sources of health care. A multistage sample first obtained information concerning the presence of children and the usual source of medical care; this information then was used to select, from households with children between the ages of six months and eleven years, a random sample stratified by type of usual source of care. A single respondent (usually the mother) was interviewed in each household and provided extensive information on the family unit and on specific family members. The completion rate was 85-90%.

The present study deals with a subset of the original sample, consisting of all persons (1) who had a regular source of health care, and (2) whose data were complete

⁷ Other estimation methods would have been more appropriate for some of these measures, e.g., Logit for dichotomous variables or Tobit for variables with

a limited range (Goldberger, 1966:248). However, selected analyses of study data as well as evidence from other studies suggest that the major substantive results obtained using OLS are similar or identical to those based on one of the other methods (Jackson, 1970; Richardson, 1971:141; Dutton, 1976:440).

from both the household interview and their regular providers. The subset was roughly half the size of the original sample, due primarily to missing data from regular providers that had not been included in the provider survey. (Fewer than 10% of persons in the original sample were eliminated because they did not name a regular source of care.) The subsample included 681 families, containing 1,623 adults (12-years-old or more) and 1,435 children. It closely resembled the original sample on key factors such as socioeconomic status and illness levels.

A single usual provider was identified for each family member, based on the regular sources of care and utilization reported in the household interview.⁸ Information on these providers came from a separate survey of all regular sources of children's medical care mentioned in the household interviews. Providers included medical institutions as well as particular professionals within institutions. The response rate was 80–85%. The subsample of providers analyzed in the present study included 75 solo practitioners, 16 fee-for-service (FFS) group practices, three prepaid group clinics (all belonging to a single prepaid plan), and five hospital outpatient departments or emergency rooms (OPD/ERs).

Since the study was conducted in Washington, D.C., many (about 90%) of the study families, as well as an unusual number of the providers, were black. The regression analyses distinguished the effects of race from those of other socioeconomic and attitudinal characteristics which might affect utilization. However, the results may still reflect some patterns which are more characteristic of black than of white populations. With respect to income distribution and utilization rates, the sample was quite similar to U.S. urban averages (NCHS, 1972). Providers repre-

sented the five health care systems also appeared to be quite characteristic, with the exception of racial distribution. Nevertheless, since the findings are based on the particular systems and patient groups studied, they may or may not be generalizable to other population groups and providers.

FINDINGS

Do the Poor Use Fewer Discretionary Health Services?

When age, sex, and illness levels are taken into account, are there income trends in discretionary use? Estimates from the initial regression model, shown in Table 1, address this question. They indicate a strong and significant positive relationship between income and every measure of discretionary use. Table 2 shows the differences in estimated use rates between the highest and lowest income groups in the Stage 1 model, with the other factors in the equation (age, sex, and illness) held constant. These figures give some idea of the magnitude of the income-related differentials which remain after age, sex, and illness are taken into account; for instance, the likelihood of a high-income child with a given illness level seeing a doctor during the six month period was more than 50% greater than that of a comparably ill low income child. Low use by the poor, according to these figures, is still very much with us.

Is it income alone that accounts for these trends or some other factors associated with poverty, such as education or occupation? Lefcowitz (1973) has argued, for example, that the association between income and use found in national statistics depends primarily on lifestyle measures such as education. Estimates from the Stage II model, also shown in Table 1, deal with this issue. They indicate that, even accounting for socioeconomic factors plus measures of family structure (in addition to demographic and illness measures), people with lower incomes were still less likely to use preventive services and to initiate symptomatic care. While the income trends in Stage II were much weaker than those in the initial model, they were (with only one exception) uniformly posi-

⁸ When more than one regular source was listed, the provider seen most often during the last six months was defined as the usual provider. While other providers may also have been seen, use of the usual providers, as defined, generally constituted a large proportion of total use; for example, among persons with visits during the six months prior to the survey, more than 95% saw their usual provider and 80% saw only their usual provider.

Table 1. Comparison of the Relation between Income and Use in Stage I and Stage II Models, before and after the Addition of Other Socioeconomic Measures

Measures of Discretionary Use	Standardized Beta Coefficients	
	Stage I (including only age, sex, and illness)	Stage II (including socioeconomic and family variables)
PREVENTIVE USE		
Number of <i>respondent's</i> checkups (n=679)	.249***	.073*
Frequency of <i>children's</i> checkups (n=1,206)	.199***	.073*
INITIATION OF CARE		
Probability of any visits to usual provider in last six months by <i>person</i> (n=3,033)	.175***	.046*
Probability of any visits to usual provider in last six months by <i>child</i> (n=1,206)	.227***	.124***
Probability that <i>child</i> ever had a hearing test or ear examination (n=1,206)	.108***	-.019
Probability that <i>child</i> ever had a vision test or eye examination (n=959)	.148***	.067*
Probability that <i>child</i> saw doctor for past ear problem (n=371)	.144***	.102*

*** indicates $p < .01$ ($t \geq 2.58$).

** indicates $p < .05$ ($1.96 \leq t < 2.58$).

* indicates $p < .20$ ($1.28 \leq t < 1.96$).

n indicates number of cases analyzed in each estimation equation.

tive and at least somewhat statistically significant. Thus, the association between income and use appears to be independent of certain objective lifestyle factors. The following sections discuss the relative validity of the three alternative explanations for this association which were described in the introduction.

Why Do the Poor Use Less Care? Assessing the Three Types of Explanations

(1) *The financial coverage explanation.* Since public assistance and private health insurance both offered financial coverage, albeit to different income groups, both were included in testing the financial coverage explanation of low use rates. Recall that according to this explanation, the poor use less care primarily because they cannot afford the costs and are not insured against them. Private insurance coverage is certainly less common among the poor in general (Schultze et al., 1972:217), and in the present study; the proportion of study families with private health insurance⁹ varied directly with in-

come, from less than 20% of the poor with some private coverage to more than 95% of upper income families. But more than half of the poverty level families received public assistance. Nonetheless, many families were still left without adequate insurance, public or private, and even families covered by public assistance may have experienced other (noneconomic) barriers in seeking care.¹⁰

The Stage III regression results (Table 3) suggest that, despite the possible inadequacies of public assistance, varying levels of financial coverage had little to do with the initial positive income trends in use. The changes in the income estimates when the two financial coverage measures were added indicate that these measures

coverage (including Medicare but excluding Medicaid). No information was obtained about the extent of financial protection or the range of services covered.

¹⁰ A serious problem in the Medicaid program is the large number of low-income families that are either not eligible for coverage or fail to enroll (Blake, 1973; Davis and Reynolds, 1976). But even those who are covered tend to have lower use rates *relative to disability levels* than do upper income groups (Aday, 1975; Davis, 1976). As Blake (1973:18) succinctly put it, "Medicaid is a case study of the failure of financial insurance alone to guarantee adequate health care."

⁹ Unfortunately, study data on private insurance coverage indicate only whether the family had

Table 2. Estimated Differences in Use between Highest and Lowest Income Groups Due to Income Alone: Comparisons from Models Successively Taking Financial Access, Attitudes, and Systems into Account

Measures of Discretionary Use	Estimated Differences in Use of Highest and Lowest Income Levels Due to Income				
	Stage I (including age, sex, and illness)	Stage II (adding family and socio- economic factors)	Stage III (adding financial access)	Stage IV (adding health-related attitudes)	Stage V (adding health care systems)
PREVENTIVE USE					
Number of <i>respondent's</i> checkups (n=679)	.959***	.279*	.274*	.189	.071
Frequency of <i>children's</i> checkups (n=1,206)	1.33***	.489*	.445*	.243	.195
INITIATION OF CARE					
Probability of any visits to usual provider in last six months by <i>person</i> (n=3,033)	.466***	.124*	.134**	.119*	.041
Probability of any visits to usual provider in last six months by <i>child</i> (n=1,206)	.522***	.323***	.351***	.303***	.302***
Probability that <i>child</i> ever had a hearing test or ear examination (n=1,206)	.281***	-.049	-.043	-.073	-.122
Probability that <i>child</i> ever had a vision test or eye examination (n=959)	.365***	.166*	.161*	.131	.046
Probability that <i>child</i> saw doctor for past ear problem (n=371)	.300***	.245*	.197	.175	.158

*** indicates $p < .01$ ($t \geq 2.58$).

** indicates $p < .05$ ($1.96 \leq t < 2.58$).

* indicates $p < .20$ ($1.28 \leq t < 1.96$).

n indicates number of cases analyzed in each estimation equation.

Table 3. Comparison of the Relation between Income and Use in Stage III, Stage IV, and Stage V Models, as Financial Access, Health-Related Attitudes, and System Type Variables Are Added

Measures of Discretionary Use	Standardized Beta Coefficients		
	(Changes in beta from previous stage in parentheses)		
	Stage III adding financial coverage (Change: III-II)	Stage IV adding attitudes (Change: IV-III)	Stage V adding system types (Change: V-IV)
PREVENTIVE USE			
Number of <i>respondent's</i> checkups (n=679)	.071* (-.002)	.049 (-.022)	.018 (-.031)
Frequency of <i>children's</i> checkups (n=1,206)	.066* (-.007)	.036 (-.030)	.029 (-.007)
INITIATION OF CARE			
Probability of any visits to usual provider in last six months by <i>person</i> (n=3,033)	.050** (.004)	.045* (-.005)	.015 (-.030)
Probability of any visits to usual provider in last six months by <i>child</i> (n=1,206)	.135*** (.011)	.116*** (-.019)	.115*** (-.001)
Probability that <i>child</i> ever had a hearing test or ear examination (n=1,206)	-.017 (.002)	-.028 (-.011)	-.047 (-.019)
Probability that <i>child</i> ever had a vision test or eye examination (n=959)	.065* (-.002)	.053 (-.012)	.018 (-.035)
Probability that <i>child</i> saw doctor for past ear problem (n=371)	.094 (-.008)	.084 (-.010)	.075 (-.001)

*** indicates $p < .01$ ($t \geq 2.58$).** indicates $p < .05$ ($1.96 \leq t < 2.58$).* indicates $p < .20$ ($1.28 \leq t < 1.96$).

n indicates number of cases analyzed in each estimation equation.

had a very minor impact on the initial estimates. The largest changes occurred in the probability of initiating care, where the positive income trends were strengthened (both in magnitude and significance) by adding financial coverage measures to the equation. This suggests that, although both forms of coverage probably promoted use, public assistance had a greater positive effect on the poor than did private insurance on the nonpoor.

There are several probable reasons for the stronger impact of public assistance. First, unlike many private insurance policies, public assistance covered most physician services (U.S. House of Representatives, 1976). Second, with the exception of prepaid health plans, even when private insurance does cover physician services, it generally appears to have minimal impact on initiation of care (Phelps,

1975). Medicaid coverage, on the other hand, is clearly associated with higher rates of use (Aday, 1975). Indeed, one would expect financial coverage to be more essential to the poor than to those more affluent.

The importance of Medicaid coverage is also evident in the Stage III estimates shown in Table 2. Before taking account of financial coverage variables, children in the highest income families were about 32% more likely to see their doctor during the six-month period than were children in families below the poverty level. When financial assistance was added, the estimated difference in these likelihoods was increased to 35%. In other words, the initial income-use relation was (mis)estimated on the basis of use rates among the poor which were inflated by public assistance. When public assistance was in-

cluded in the model, its positive effect could be properly assigned, thereby deflating the apparent use rates of the poor and producing a more strongly *positive* estimate of the income use relation. This is illustrated in Figure 2.

In short, lack of financial coverage does not appear to explain low use rates among the poor in our study; rather, positive income trends in use exist *despite* the positive effects of public assistance on the low income families covered.

(2) *The culture of poverty explanation.* According to proponents of the culture of poverty viewpoint, low use rates are fundamentally due to attitudes and preferences characteristic of the poor. A wide variety of such attitudes have been studied, with mixed findings concerning not only the importance of such measures in explaining utilization, but also their predominance among the poor (Shaw, 1970; Bice, 1971). Much confusion concerning the role of various attitudes can be traced, in fact, to the failure to specify whether they are viewed as independent determinants of use or rather as mediating links between other objective characteristics—such as socioeconomic status—and use.

In the present study, culturally-based values and practices are represented by five health-related attitude measures:¹¹ (1) belief in preventive checkups, an index based on four questions concerning the importance of seeing a doctor regularly even without illness symptoms; (2) tendency to consult a physician, an index of responses concerning the likelihood of consulting a physician for a set of common symptoms; (3) professional health orientation, an index of thermometer ownership and use; (4) the salience of health, an index of the frequency of discussion of health matters; and (5) social alienation, an index based on questions developed by Srole (1956) to assess general social

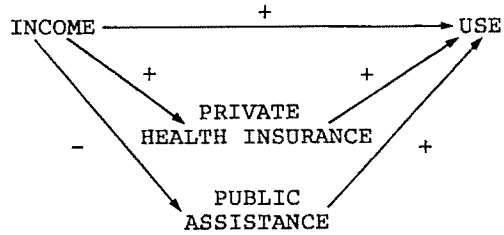


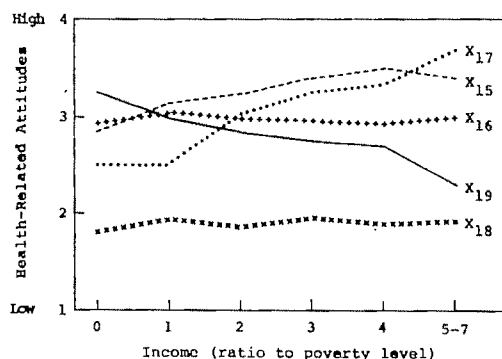
Figure 2. General Path Model Showing the Intervening Role of Financial Coverage Measures

anomie. These five measures were added to the regression model in order to evaluate the validity of the culture of poverty explanation of low use rates.

Belief in preventive checkups approximates one of the measures in Rosenstock's (1966) very influential health belief model—belief in the effectiveness of preventive action in reducing the threat posed by disease. This model has been the subject of considerable research (Becker et al., 1977a); many findings support the importance of belief in effectiveness, in particular. The tendency to consult a physician measure reflects differences in defining symptoms as illness, as well as differences in the response to these symptoms (Battistella, 1968), both of which may vary by social class (Mechanic and Volkart, 1961; Hetherington and Hopkins, 1969; Rosenstock, 1975). The indicator of professional health orientation (based on thermometer usage) represents a cosmopolitan attitude toward health care (Suchman, 1966) reinforced, probably, by previous contacts with the professional health care system (Rosenstock, 1975). It has also been alleged that health is less salient to the poor—that they are more likely to put up with illness symptoms (Koos, 1954), or not to define them as illness (Zola, 1966; Zborowski, 1958), and consequently, not to seek care. Finally, the index of social alienation represented the passivity and resignation ascribed to the culture of poverty, which have often been cited as a cause of low utilization, especially of preventive services (Moody and Gray, 1972; Bullough, 1972; Hyman, 1970; Becker et al., 1977a).

As mentioned, findings from previous research differ concerning the relationship of various attitudes to economic status. In

¹¹ These measures, listed in Appendix C, were obtained from the respondents in the household interviews, who were generally the mothers. Their attitudes are thus directly relevant to their own and their children's utilization, and are likely to be at least somewhat correlated with the attitudes of other adult family members.



Note: All attitude values have been transposed to a 1-4 scale.

- X₁₅ Belief in preventive checkups
- X₁₆ Tendency to consult a physician
- X₁₇ Professional health orientation: index of thermometer use
- X₁₈ Salience of health: frequency of communication about health matters
- X₁₉ Social alienation index

Figure 3. Variation in Health-Related Attitudes by Income Level

the study sample, three of the attitudes varied by income level and two did not (see Figure 3). The indices of tendency to consult a physician and salience of health showed almost no variation, in contrast to many previous findings (McKinlay and Dutton, 1974). The three attitudes that did vary were: belief in preventive checkups, professional health orientation (both of which increased with income), and social alienation (which decreased). Of course, these differences may not reflect cultural variation so much as realistic adaptation to economic circumstances; preventive care may well be less important, for example, than paying the rent, and purchasing a thermometer may be viewed as an unaffordable luxury.

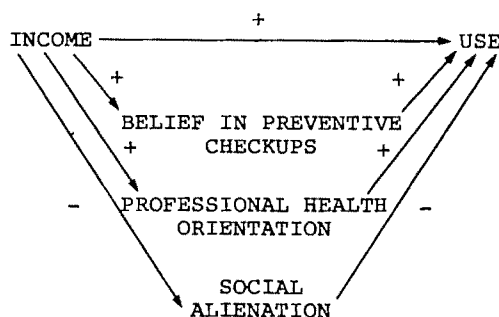


Figure 4. General Path Model Showing the Intervening Role of Health-Related Attitudes

Attitudes can mediate the income-use relationship either through an attitude linked positively to both income and use (such as belief in checkups) or, alternatively, through an attitude linked negatively to both income and use (such as alienation). This is illustrated in Figure 4. In both cases, omitting the mediating attitude from the model causes the direct income-use relation to be overestimated—that is, too strong a positive relation is estimated.

To what extent do these attitudes account for the remaining income trends in use? In light of the earlier discussion and recent research (Kirscht et al., 1976; Becker et al., 1977b), the culture of poverty explanation was modified somewhat so as not to require attitudes to predominate over financial access. Instead, financial coverage differences were accounted for before examining the mediating role of attitudes, simply by adding (in Stage IV) the five attitude measures to the preceding model (which included financial coverage plus the other family and personal characteristics). The estimates obtained are shown in Table 3. In all cases, the previous estimates for income were reduced when attitudes were added.

These changes suggest that attitudes played a rather important part in explaining the income trends in use, especially preventive use.¹² Although in both preventive use equations the estimates for income remained positive, neither was (even weakly) statistically significant.

Income estimates were also reduced in the initiation of care equations, but not eliminated; in fact, both probability of visits estimates were still somewhat significant ($p < .01$ for children and $p < .20$ for all family members). In addition, Table 2 shows that the probability of poor children seeking medical care was still about 30% less than that of the most affluent children, even after illness level, family char-

¹² Recent data suggest that the relation of attitudes to economic status is complex, and varies according to other characteristics such as race and age (Kravits, 1975). It appears, in fact, that attitudes may be more closely correlated with utilization among blacks than among whites. If so, findings from the present study may overestimate the importance of attitudes for a general population sample.

acteristics, health attitudes, and financial coverage were all taken into account.

In sum, while the financial coverage and culture of poverty explanations together appear to provide a sufficient explanation for low rates of preventive use by the poor, they failed to account fully for income differentials in seeking symptomatic care.

(3) *The system barriers explanation.* In this study, as in most reported data, use of different health care systems tended to be divided along economic lines. About three-quarters of the families at or below the poverty level reported either a hospital outpatient department or emergency room (OPD/ER) or a public clinic as their usual source of primary health care; roughly a third of the total patient clientele in each of these systems was at or below the poverty level. In contrast, most upper income families used traditional fee-for-service providers or the prepaid group practice.

As noted, the clinic care given to the poor in many institutions has often been criticized for its dehumanizing ambience and structural problems. The deterrent effect of specific features of such systems have been demonstrated (Dutton, 1977; Aday and Andersen, 1975). Moreover, these effects appear to account for the remaining association between income and initiation of care. When variables representing the usual system of care were added to the model, the positive income trends were greatly reduced (see Table 3, Stage V). In fact, the average reduction in these estimates was larger than in any of the preceding Stages.¹³ All of the income estimates (except in the initiation of children's care equation) are now insignifi-

cant. Table 2 again shows how much the income differentials have narrowed. After systems were added, the estimated probability of members' visits was only 4% greater for the most affluent families than for the poorest families (down from a 12% differential before systems were accounted for).

This suggests, in short, that use rates are low among the poor in part *because of inadequacies in the health care systems they use*. Some of these inadequacies are described below, along with estimates of selected system effects on use based on results from the full model.

Comparing the Direct Effects of Financial Coverage, Attitudes, and Health Care Systems

As successive groups of variables were added to the model, explanatory power originally attributed to income was reallocated to other variables. The direct effects on use estimated for each group of variables in the full model indicate which had the strongest independent effects on use, and also which were the most important mediators of the original income-use relations.

Table 4 compares the estimates for income in the full (Stage V) model with the strongest estimates for variables in each of the three groups of explanatory factors.¹⁴ As is evident, the income estimates generally were weak relative to those of the other variables. (The only exception was, once again, the probability of children's visits, where the income effect was very strong.) For most use measures, then, income-related factors, particularly the type of health care system used, play a more important role than income itself in determining patterns of discretionary use.

System effects, measured relative to

¹³ Moreover, because income, financial coverage, attitudes, and health care system types were all highly correlated, the order in which these variables are added to the model affects their apparent impact on income estimates at each stage. The impact of variables entered early includes covariation shared with variables not yet entered. Thus, comparison of changes in the income coefficients after each group of variables was added overrepresents the impact of variables added first, and underrepresents the impact of those added last, namely, the health care system variables. This procedure was used, nonetheless, in order to illustrate the insufficiency, even in combination, of alternative explanations of the income-use relation.

¹⁴ The proportion of total variation in each use measure explained by all of the variables in the Stage V model (R^2) ranged from .112 for consulting a doctor about a child's ear problems, to .254 for probability of a child hearing test. Thus, despite the wide range of factors included in this model, much of the variation in use rates remains unexplained. These figures are similar to or higher than those in most comparable analyses, however, and they do not diminish the credibility of significant relationships estimated for the variables analyzed.

Table 4. Estimated Direct Effects on Use from Stage V Model Comparing Measures of Financial Coverage, Attitudes, and Health Care System Types *

Measures of Discretionary Use	Financial Coverage		Attitudes	Health Care Systems	
	(Stage V) Income	Public Assistance	Belief in Preventive Checkups	OPD/ER vs. Solo MD	PGP vs. Solo MD
PREVENTIVE USE					
Number of <i>respondent's</i> checkups (n=679)	.018 (.34)	.052 (1.05)	.136*** (3.55)	-.111** (-.245)	.040 (.96)
Frequency of <i>children's</i> checkups (n=1,206)	.029 (.74)	.023 (.62)	.111*** (4.06)	-.102*** (-3.07)	.147*** (5.05)
INITIATION OF CARE					
Probability of any visits to usual provider in last six months by <i>person</i> (n=3,033)	.015 (.60)	.071*** (3.02)	.037** (2.10)	-.056** (-2.64)	.098*** (5.21)
Probability of any visits to usual provider in last six months by <i>child</i> (n=1,206)	.116*** (2.87)	.063* (1.62)	.079*** (2.84)	.029 (.85)	.096*** (3.24)
Probability that <i>child</i> ever had a hearing test or ear examination (n=1,206)	-.047 (1.17)	.082** (2.14)	.096*** (3.46)	-.003 (-.077)	.151*** (5.10)
Probability that <i>child</i> ever had a vision test or eye examination (n=959)	.018 (.38)	0.080* (1.80)	.097*** (2.98)	-.058* (-1.48)	.138*** (4.05)
Probability that <i>child</i> saw doctor for past ear problem (n=371)	.075 (.94)	.026 (.35)	.111* (1.92)	.086* (1.30)	.148** (2.47)

* Standardized beta coefficients; t statistics in parentheses.

*** indicates $p < .01$ ($t \geq 2.58$).** indicates $p < .05$ ($1.96 \leq t < 2.58$).* indicates $p < .20$ ($1.28 \leq t < 1.96$).

n indicates number of cases analyzed in each estimation equation.

solo practice, were estimated, with variation in access, attitudes, and other patient characteristics taken into account. Thus, positive estimates mean that the system promoted a given type of use relative to solo practice, while negative estimates indicate a relative deterrent effect. As seen in Table 4, the OPD/ERs generally had the strongest *negative* impact on discretionary use, while the prepaid group practice (PGP) consistently had the most *positive* impact.

These disparities stem at least in part from major structural differences among the various systems (Dutton, 1978). Consider the following comparisons of OPD/ERs and the prepaid group. Many OPD/ERs probably did not even offer preventive health examinations. In contrast, pre-

vention is the theoretical cornerstone of prepaid group practice. Second, average charges at OPD/ERs were higher than those in solo practice, while for prepaid group patients services were free of charge after the initial enrollment sum. Waiting times in the OPD/ERs averaged almost an hour, in contrast to 20 minutes in the prepaid group. And relationships between patients and physicians were much poorer in the OPD/ERs than in the prepaid group; not one of the OPD/ER physicians reported being "satisfied with most or all" of their patients, in contrast to virtually all (97%) of the prepaid group physicians reporting such satisfaction. Barriers such as these in the OPD/ERs apparently discouraged patients—mainly the poor—from seeking care, *above and*

beyond the deterrent effects of inadequate financial coverage and negative attitudes toward professional health care.

Finally, it is worth mentioning how some of the Stage V estimates shown in Table 4 differed from those of Stages II-IV, which did not include system types. First, estimates of the independent association between attitudes and use were somewhat *weakened* when system differences were taken into account. Since people with use-promoting attitudes tended to use use-promoting systems, the earlier estimates confounded system effects with attitudinal effects. (Of course, it is difficult to separate cause from effect here; very likely, it is a mutually reinforcing process.) Conversely, estimates of the impact of public assistance on use were *strengthened* substantially when system types were added to the model. Prior to controlling for system types, the positive effects of financial coverage were offset partially by the negative effects of the health care systems typically used by persons with public assistance (e.g., OPD/ERs). The two effects were confounded when system differences were not taken into account, and the importance of public assistance, therefore, was underestimated.

CONCLUSIONS

This paper has presented an empirical assessment of alternative explanations of low use of discretionary health care services by the poor. Two of the most prominent explanations—one based on inadequate financial coverage and the other on cultural attributes of the poor—appear to be insufficient in explaining income differentials in initiating symptomatic care. These differentials are also attributable to inadequacies in the health care systems often used by the poor, notably hospital outpatient departments and emergency rooms. Particularly striking were the contrasts between such systems and the pre-paid group practice, a newer mode of health care delivery serving primarily middle- and upper-class patients.

Such contrasts raise serious questions about the validity of findings from previous utilization studies which have ignored

differences among the systems being used. These findings are likely to be biased, in ways which are suggested, by the changes in estimates reported in the present study after system types were added to the model. Recall that since the poor generally used systems which inhibited use, failure to account for system differences resulted in underestimating the importance of factors which promoted use among the poor, such as public assistance. On the other hand, the importance of various personal traits of the poor, such as attitudes, was overemphasized. In both instances, because system differences were ignored, the findings were misleading and the policy implications correspondingly distorted. Too frequently, as Gibson (1971:637) has said, "It is the individual patient (or potential patient) that is 'blamed' and the system ignored in the search for causes and rationales for change."

Furthermore, such a focus has diverted attention away from inadequacies in the health care systems used by the poor. Explaining low use rates in terms of the inadequacies of these systems has obvious implications for policies aimed at improving access (Goering and Coe, 1970). Such policies inevitably would require substantial redistribution of health resources, and research with this orientation may have had lower funding priority than research directed at alternative methods of improving access. The policy implications of research on the culture of poverty probably seemed more manageable—and less threatening—than the potential implications of research on the culture of medicine. Despite the lack of empirical research, of course, there has been mounting criticism of this medical subculture, with its emphasis on high-technology, mass-produced, super-specialty care. Certainly, it has failed to meet the health needs of the poor in many large institutions.

The frequent failure to consider the inadequacies of health care systems used by the poor in investigative designs may also be due to limited contact by researchers with such systems—one wonders, for example, how many social scientists have actually received care at a hospital out-

patient department. Lack of contact may be further compounded by disciplinary frameworks that constrain explanatory strategies.

In any event, there is little excuse for continuing to ignore the important effects that different health care systems have on patient utilization. Findings from the present study add to the growing body of evidence that low use rates among the poor are explained in part by the systems

of care they use. These findings suggest that neither improved financial access nor health education efforts alone will eliminate current income differentials in use, unless accompanied by structural improvements in existing delivery systems. Fundamental changes in the organization and distribution of care must occur, if equitable patterns of use are to be more than health policy rhetoric.

APPENDIX A

Dependent Variables: Measures of Discretionary Ambulatory Care Use

	Coding	Number of Cases	Mean Values
PREVENTIVE USE			
Y ₁ Number of preventive health examinations reported by <i>respondent</i>	0,1,2+	679	1.5
Y ₂ Frequency with which <i>children</i> are taken for preventive health examinations	1=never 2=when needed 3=every year 4=every six months	1206	2.6
INITIATION OF CARE			
Y ₁ Did <i>person</i> have any visits to usual provider in last six months? ^a	1=yes 0=no	3303	.43
Y ₂ Did <i>child</i> have any visits to usual provider in last six months?	1=yes 0=no	1206	.57
Y ₃ Has <i>child</i> ever had a (nonschool) hearing test or ear examination?	1=yes 0=no	1206	.43
Y ₄ Has <i>child</i> ever had a (nonschool) vision test or eye examination? ^b	1=yes 0=no	959	.29
Y ₇ Did child see a doctor for a past ear problem? ^c	1=yes 0=no	371	.75

^a Includes both children and adults.

^b Asked only of children aged 4-11 years.

^c Asked only of children who were reported in the household interview to have had ear problems.

APPENDIX B

Explanatory Variables in the Stage I and Stage II Models of Use

	Coding	Mean Values
REPORTED INDIVIDUAL ILLNESS MEASURES ^a		
X ₁ Child's illness index (a weighted proportion of the following: prematurity, ear problem, hearing test failure, eye problems, vision test failure, sinus trouble, hay fever, asthma, eczema, hives, chronic food allergy, more than 3 colds/year) ^b	continuous index, from 0=no problem, to 1=all problems	0.11
X ₁ Child's ear illness index (based on ear problems and hearing test failure) ^b	0=no problem, to 1=both	0.25
X ₁ Child's eye illness index (based on eye problems and vision test failure) ^b	0=no problem, to 1=both	0.22
X ₁ Does person have a current medical or health problem? ^a	1=yes, 0=no	0.10
DEMOGRAPHIC FACTORS		
X ₂ Individual's age	number of years	17.4
X ₃ Individual's sex	1=female, 0=male	0.53
FAMILY STRUCTURE		
X ₄ Age of head-of-house	number of years	34.3
X ₅ Sex of head-of-house	1=female, 0=male	0.31
X ₆ Family size	number of members	4.5
X ₇ First-born child ^a	1=first born 0=subsequent child	0.36
SOCIOECONOMIC FACTORS		
X ₈ Family income (ratio to poverty level)	0=below poverty, to 7=seven times	2.1
X ₉ Occupation of head-of-house	1=blue-collar 0=white collar	0.56
X ₁₀ Education of respondent (years of school)	1=8, 2=9-11, 3=12, 4=13-15, 5=16+	2.96
X ₁₁ Race	1=black, 0=white	0.93
X ₁₂ Length of residence in D.C.	number of years	19.0

^a All illness measures are designated X₁ since only one appears in each use equation; e.g., ear illness in ear exam equation (Y₄), general illness index in any visits equation (Y₅), etc.

^b Appears only in children's initiation of care equations (Y₆-Y₇).

^c Appears only in adult initiation of care equations (Y₈).

^d Appears in all children's use equations.

APPENDIX C
Explanatory Variables in Stages III, IV, and V

STAGE III: Financial Coverage		Coding	Mean Values
X ₁₃	Private health insurance	1=yes, 0=no	0.71
X ₁₄	Public assistance	1=yes, 0=no	0.271
STAGE IV: Health-Related Attitudes of Respondent			
X ₁₅	Belief in preventive checkups (index based on three questions: How important is it to go to a doctor when well? As long as you feel well, is there no need to go to a doctor? Although not sick, should you see a doctor once a year?)	1=low, to 4=high	3.3
X ₁₆	Tendency to consult a physician (index, averaging responses to seven questions on illness symptoms and how likely respondent would be to consult doctor for each)	1=very unlikely, to 4=very likely	3.0
X ₁₇	Professional health orientation (index based on thermometer ownership and use)	1=don't own thermometer 2=own, but don't use regularly 3=own and use before calling M.D.	2.3
X ₁₈	Salience of health (based on frequency of discussion about health matters)	1=never, to 4=very frequently	1.9
X ₁₉	Social alienation (index of four anomie measures, based on Srole (1956): Roughly, public officials are not interested in the average person; people don't care what happens to others; might as well live for today; parents today are too easy with their children)	0=not alienated, to 1=very alienated	0.62
STAGE V: Usual Type of Health Care System			
—	Fee-for-service solo practice	1=yes, 0=no	0.27
X ₂₀	Fee-for-service group practice	1=yes, 0=no	0.10
X ₂₁	Prepaid group practice	1=yes, 0=no	0.12
X ₂₂	Public clinic	1=yes, 0=no	0.25
X ₂₃	Hospital outpatient department or emergency room	1=yes, 0=no	0.26

REFERENCES

- Aday, L. A.
1975 "Economic and noneconomic barriers to the use of needed medical services." *Medical Care* 13:447-56.
- Aday, L. A. and R. Andersen
1975 *Development of Indices of Access to Medical Care*. Ann Arbor: Health Administration Press.
- Andersen, R.
1975 "Health service distribution and equity." Pp. 9-30 in R. Andersen, J. Kravits, O. W. Andersen (eds.), *Equity in Health Services*. Cambridge, Ma.: Ballinger.
- Andersen, R. and J. F. Newman
1973 "Societal and individual determinants of medical care utilization in the united states." *Milbank Memorial Fund Quarterly* 51:95-124.
- Battistella, R. M.
1968 "Limitations in use of the concept of psychological readiness to initiate health care." *Medical Care* 6:308-19.
- Becker, M. H. and L. A. Maiman
1975 "Sociobehavioral determinants of compliance with health and medical care recommendations." *Medical Care* 13:10-24.
- Becker, M. H., D. P. Haefner, S. V. Kasl, J. P. Kirscht, L. A. Maiman, I. M. Rosenstock
1977 "Selected psychosocial models and correlates of individual health-related behaviors." *Medical Care* 15:27-46.
- Becker, M. H., C. A. Nathanson, R. H. Drachman, and J. P. Kirscht
1977b "Mothers' health beliefs and children's clinic visits: a prospective study." *Journal of Community Health* 3:125-35.
- Berkanovic, E. and L. G. Reeder
1974 "Can money buy the appropriate use of

- services? some notes on the meaning of utilization data." *Journal of Health and Social Behavior* 15:93-9.
- Bice, T. W.
1971 *Medical Care for the Disadvantaged*. Baltimore: Johns Hopkins University Department of Medical Care and Hospitals.
- Blake, E.
1973 "Medicaid: the fading of a dream." *Health PAC Bulletin* 51:13-9.
- Blauner, R.
1964 *Alienation and Freedom*. Chicago: University of Chicago Press.
- Bryant, J. H.
1977 "Principles of justice as a basis for conceptualizing a health care system." *International Journal of Health Services Research* 7:707-19.
- Bullough, B.
1972 "Poverty, ethnic identity and preventive health care." *Journal of Health and Social Behavior* 13:347-59.
- Davis, K.
1976 "Medicaid payments and utilization of medical services by the poor." *Inquiry* 13:122-35.
- Davis, K. and R. Reynolds
1976 "The impact of medicare and medicaid on access to medical care." Pp. 391-435 in R. N. Rosett (ed.), *The Role of Health Insurance in the Health Services Sector*. New York: National Bureau for Economic Research.
- Durkheim, E.
1933 *Suicide*. Glencoe: Free Press.
- Dutton, D. B.
1976 *A Causal Model of the Use of Health Services: The Role of the Delivery System*. Ph.D. dissertation, Massachusetts Institute of Technology.
1977 "Organizational and professional characteristics of health care settings and their impact on patients' utilization." Paper presented at the 105th annual meeting of the American Public Health Association, Washington, D.C.
1978 "Patterns of ambulatory health care in five different delivery systems." *Medical Care*. In press.
- Freidson, E.
1974 *Professional Dominance: The Social Structure of Medical Care*. Chicago: Aldine.
- Gibson, G.
1971 "Explanatory models and strategies for social change in health care behavior." *Social Science and Medicine* 6:636.
- Goering, J. and R. Coe
1970 "Cultural versus situational explanations of the medical behavior of the poor." *Social Science Quarterly* 51:309.
- Goldberg, V. P.
1976 "Some emerging problems of prepaid health plans in the medical system." *Policy Analysis* 2:55-68.
- Goldberger, A. S.
1966 *Econometric Theory*. New York: Wiley.
- Hetherington, R. W. and C. E. Hopkins
1969 "Symptom sensitivity: its social and cultural correlates." *Health Services Research* 4:63-75.
- Hyman, M. D.
1970 "Some links between economic status and untreated illness." *Social Science and Medicine* 4:387-99.
- Illich, I.
1976 *Medical Nemesis*. New York: Random House.
- Jackson, J. E.
1970 "Ordinal data, ordinary least squares, and senate voting." Paper presented at the Conference in Structural Equation Models, Madison.
- Kadushin, C.
1964 "Social class and the experience of ill health." *Sociological Inquiry* 34:67-80.
1967 "Social class and ill health: the need for further research. a reply to Antonovsky." *Sociological Inquiry* 37:323-32.
- Kessner, D. M., C. K. Snow, and J. Singer
1974 *Assessment of Medical Care for Children, Contrasts in Health Status*, Vol. 3. Washington, D.C.: Institute of Medicine, National Academy of Sciences.
- Kirscht, J. P., M. H. Becker and J. P. Eveland
1976 "Psychological and social factors as predictors of medical behavior." *Medical Care* 14:422-31.
- Koos, E. L.
1954 *The Health of Regionville*. New York: New York: Columbia University Press.
- Kovner, J. W., L. B. Browne, and A. I. Kisch
1969 "Income and the use of outpatient medical care by the insured." *Inquiry* 6:27-34.
- Kravits, J.
1975 "The relationship of attitudes to discretionary physician and dentist use by race and income." Pp. 73-93 in R. Andersen, J. Kravits, and O. W. Andersen (eds.), *Equity in Health Services*. Cambridge, Ma.: Ballinger.
- Kravits, J. and J. Schneider
1975 "Health care need and actual use by age, race and income." Pp. 169-87 in R. Andersen, J. Kravits, and O. W. Andersen (eds.), *Equity in Health Services*. Cambridge, Ma.: Ballinger.
- Leacock, E. G.
1971 *The Culture of Poverty: A Critique*. New York: Simon and Schuster.
- Lefcowitz, M. J.
1973 "Poverty and health: a reexamination." *Inquiry* 10:3-31.
- Levine, S., N. Scotch, and G. Vlasak
1969 "Unravelling technology-culture in public health." *American Journal of Public Health* 59:237-44.
- Lewis, O.
1966 "The culture of poverty." *Scientific American* 215:16-25.
- McKinlay, J. B.
1972 "Some approaches and problems in the study of the use of services—an overview."

- Journal of Health and Social Behavior 13:115-52.
- McKinlay, J. B. and D. B. Dutton
1974 "Social-psychological factors affecting health services utilization." Pp. 251-303 in S. T. Mushkin (ed.), *Consumer Incentives for Health Care*. New York: Prodist.
- May, J. J.
1975 "Utilization of health services and the availability of resources." Pp. 131-49 in R. Andersen, J. Kravits, and O. W. Andersen (eds.), *Equity in Health Services*. Cambridge, Ma.: Ballinger.
- Mechanic, D.
1975 "The organization of medical practice and practice orientations among physicians in prepaid and nonprepaid primary care settings." *Medical Care* 13:189-204.
- Mechanic, D. and E. Volkart
1961 "Stress, illness behavior, and the sick role." *American Sociological Review* 25:51-8.
- Monteiro, L. A.
1973 "Expense is no object: income and physician visits reconsidered." *Journal of Health and Social Behavior* 14:99-115.
- Moody, P. and R. Gray
1972 "Social class, social integration and the use of preventive health services." Pp. 240-51 in E. C. Jaco (ed.), *Patients, Physicians, and Illness*. Glencoe: Free Press.
- Morris, N. M., M. H. Hatch, and S. S. Chipman
1966 "Alienation as a deterrent to well-child supervision." *American Journal of Public Health* 56:1874-82.
- National Center for Health Statistics, U. S. Department of Health, Education, and Welfare (NCHS)
1966 *Short-Stay Hospitals, U.S., July 1963-June 1964*. Vital Health Statistics, Ser. 10, No. 30, Publication No. 1000. Washington, D.C.: U.S. Government Printing Office.
1972 *Physician Visits: Volume and Interval Since Last Visit, U.S., 1969*. Vital and Health Statistics, Ser. 10, No. 75, DHEW Publication No. (HSM) 72-1064. Washington, D.C.: U.S. Government Printing Office.
1974a *Disability Days, U.S., 1971*. Vital and Health Statistics, Ser. 10, No. 90, DHEW Publication No. (HRA)74-1517. Washington, D.C.: U.S. Government Printing Office.
1974b *Physician Visits: Volume and Interval Since Last Visit, U.S., 1971*. Vital and Health Statistics, Ser. 10, No. 97, DHEW Publication No. (HRA) 75-1524. Washington, D.C.: U.S. Government Printing Office.
- Phelps, C. E.
1975 "The effects of insurance on demand for medical care." Pp. 105-30 in R. Andersen, J. Kravits, and O. W. Andersen (eds.), *Equity in Health Services*. Cambridge, Ma.: Ballinger.
- Rainwater, L.
1968 "The lower class: health, illness and medical institutions." Pp. 259-77 in Irwin Deutscher and Elizabeth Thompson (eds.), *Among the People: Encounters with the Poor*. New York: Basic Books.
- Richardson, W. C.
1970 "Measuring the urban poor's use of physician services in response to illness episodes." *Medical Care* 8:132-42.
1971 *Ambulatory Use of Physicians' Services in Response to Illness Episodes in a Low-Income Neighborhood*. Research Series No. 29, University of Chicago Center for Health Administration.
- Riessman, C. K.
1974 "The use of health services by the poor." *Social Policy* 5:41-9.
- Roghmman, K.
1974 "Looking for the medical care crisis in utilization data." *Inquiry* 11(4):282-91.
- Rosenstock, I. M.
1966 "Why people use health services." *Milbank Memorial Fund Quarterly* 44:94.
1975 "Prevention of illness and maintenance of health." Pp. 193-223 in J. Kosa and I. K. Zola (eds.), *Poverty and Health: A Sociological Analysis*. Cambridge, Ma.: Harvard University Press.
- Salkever, D. S.
1975 "Economic class and differential access to care: comparisons among health care systems." *International Journal of Health Services* 5:373-95.
- Schultze, C. L., E. R. Fried, A. M. Rivlin
1972 *Setting National Priorities: the 1973 Budget*. Washington, D.C.: Brookings Institute.
- Seeman, M. and M. W. Evans
1962 "Alienation and learning in a hospital setting." *American Sociological Review* 27:772-82.
- Shaw, C. T.
1970 "Class characteristics of supporters and rejectors of basic health measures." *Social Science and Medicine* 4:411-5.
- Srole, L.
1956 "Social integration and certain correlates." *American Sociological Review* 21:709-16.
- Strauss, A. L.
1969 "Medical organization, medical care and lower income groups." *Social Science and Medicine* 3:143-77.
- Suchman, E. A.
1964 "Socio-medical variations among ethnic groups." *American Journal of Sociology* 70:319.
1966 "Social patterns of illness and medical care." *Journal of Health and Human Behavior* 6:2-16.
- U.S. Department of Health, Education, and Welfare
1968 *Report of the Task Force on Medicaid and Related Programs (McNerney Report): Summary*. Publication Number 0-398-052, Washington, D.C.: U.S. Government Printing Office.

- U.S. House of Representatives, Committee on Interstate and Foreign Commerce, Subcommittee on Health and Environment
 1976 Data on the Medicaid Program: Eligibility, Services, Expenditures, Fiscal Years 1966-76. Publication No. 61-722 0 (January). Washington, D.C.: U.S. Government Printing Office.
- Valentine, C. A.
 1968 Culture and Poverty: Critique and Counter-proposals. Chicago: University of Chicago Press.
- Weiss, J. E. and M. Greenlick
 1970 "Determinants of medical care utilization: the effect of social class and distance on contacts with the medical care system." Medical Care 8:456-62.
- Wilson, R. W. and E. L. White
 1977 "Changes in morbidity, disability, and utilization differentials between the poor and the nonpoor: data from the health interview survey: 1964 and 1973." Medical Care 15:636-46.
- Zborowski, M.
 1958 "Cultural components in response to pain." Pp. 256-68 in E. G. Jaco (ed.). Patients, Physicians, and Illness. Glencoe: Free Press.
- Zola, I. K.
 1966 "Culture and symptoms—an analysis of patients' presenting complaints." American Sociological Review 31:615-30.

SOCIAL FACTORS IN PSYCHIATRIC OUTCOME: TOWARD THE RESOLUTION OF INTERPRETIVE CONTROVERSIES

R. JAY TURNER

University of Western Ontario

JOHN W. GARTRELL

University of Alberta

American Sociological Review 1978, Vol. 43 (June):368-382

Some of the most persistently observed relationships within social science are those between certain social (or demographic) factors and the occurrence and outcome of psychological disorder. While it is widely assumed that there is an important theoretical message to be found in these linkages, a serious debate has developed concerning the nature of this message. This paper brings data to bear upon the question of the meaning of the well-known relationships of marital status, social class position, and work performance to psychiatric outcome as measured by time spent in hospital. The results suggest that these relationships are secondary associations deriving largely from social selection processes. More specifically, our findings indicate that marital status, social class, and probably work performance, appear to matter for outcome largely because varying statuses on these factors reflect important differences in social competence. Some of the implications of these results with respect to competing explanatory perspectives are discussed.

This paper is concerned with a part of the considerable body of research on schizophrenia that has been conducted from a generally sociological perspective. More specifically, the works and issues to be discussed are those that derive from sociological studies in psychiatry as distinguished from sociological studies of psychiatry (Dunham, 1971). What ties the studies we are concerned with together is an interest in the possible role of social factors in the course and outcome of psychological disorder.

The present study is of men called schizophrenic and focuses upon the well-known relationships of marital status, social class position, and work per-

formance to psychiatric outcome as measured by time spent in hospital.¹ These relationships are well-known both because they have been reported persistently by studies in widely differing settings and using widely differing methods, and because of the provocative and contradictory hypotheses that have been generated to explain them. Our purpose includes, but goes well beyond, the verification of

¹ It should be noted that the question of diagnostic validity is not of central concern here because the relationships at issue are observed quite regularly for general populations of men hospitalized for psychiatric disorder. However, evidence that the findings to be reported do apply to schizophrenia specifically is presented toward the end of this paper.

previously reported relationships. The primary goal is to present data that bears upon the crucial and unresolved question of how these relationships are to be interpreted.

Social Factors and Schizophrenic Outcome

Numerous studies have established a relationship between marital status and schizophrenic outcome or prognosis as measured by durations of hospitalization with single men spending the most time in hospital and married men the least time (see Norris, 1956; Sherman et al., 1964; Gittelman-Klein and Klein, 1968; Harrow et al., 1969; Rosen et al., 1971; Gove and Fain, 1970). Providing additional evidence for such a relationship, Mason and associates (1960:5) report being married to be more closely associated with early discharge than were any of the numerous clinical and demographic variables studied. Indeed, Farina et al. (1963) suggest that marital status may be a reliable predictive index of recovery. This latter point is also argued by Chapman et al. (1961) whose data revealed that marital status was highly related to Elgin Prognostic Scale scores and predicted length of hospitalization nearly as well.

Perhaps the best known and most provocative of all findings on schizophrenic prognosis is that lower-class patients appear to have poorer prospects for recovery, at least as indexed by treatment history. Thus Hollingshead and Redlich (1958) observe that lower-class patients tend to have longer continuous durations of stay in hospital, a more prolonged period of treatment, and a higher rate of reentry into treatment. Working in Great Britain, Brooke (1957) studied a large cohort of first-admission male schizophrenics and observed a clear social class gradient in terms of duration of hospital stay. Similarly, both Brown (1959) and Cooper (1961) provide evidence that patients from lower-class backgrounds are more likely to become long-stay patients (see also Hardt and Feinhandler, 1959; and Myers and Bean, 1968). While not every study has reported such findings (see Wing et al., 1959; Carstairs et al.,

1955), the great weight of available evidence indicates that, at least for men, social class position and time spent in hospital are reliably associated.

Although the evidence concerning work performance is somewhat less direct than that for marital status and social class, it is nevertheless quite compelling. This evidence is from two sources: (1) a body of research relating premorbid social functioning with length of hospitalization in which work performance represents a major dimension of social functioning (see Zigler and Phillips, 1961; Phillips, 1968); and (2) an array of studies that consistently show poor performance among exhospital patients to be significantly related to the likelihood of being rehospitalized (Cohen, 1955; Maisel, 1967; Brown et al., 1958; Brown, 1959; Kris, 1963).

While few would challenge the empirical reality of the relationships of marital status, social class and work performance to the outcome of disorder as typically indexed, there has been considerable debate about how they should be interpreted. Even among researchers who generally assumed that there is an important theoretical and/or practical message to be found in these linkages, the difficulty has been with determining the nature of this message.

Causal Association or Selection Artifact?

The most fundamental interpretive issue involves the question of whether these persistently observed relationships can be said to reflect a theoretically interesting connection. Is there something about social class position, marital status and work performance that have direct consequences for time spent in hospital, or are these relationships simply artifacts of social selection processes?

Proponents of the causal association position argue that different statuses place the individual in social systems that vary widely in the kind and degree of influence they exert. Factors in the social milieu, such as those indexed by these variables, are seen as influencing individuals and/or the treatment systems with which they may be involved, and largely determining

whether, when, and how long such involvement will occur. Two perspectives are consistent with this general hypothesis. One emphasizes that these and associated factors directly influence treatment systems. The other contends that such variables influence course and outcome by affecting the level of disorder or disability of the individual involved.

The first of these types is associated with societal reaction or labelling theory. In this perspective the central contingencies governing both entrance to and exit from the role of mental patient lie outside the patient's condition (Scheff, 1963; 1966). Primary among these contingencies are social resources or social power which are seen as crucial to the avoidance of labelling and, failing that, to the avoidance of hospitalization and, failing that, to the avoidance of lengthy hospitalization. Research relating to this perspective has commonly assumed that measurements of class, marital status and work performance represent reasonable operations for indexing social resources or social power (Rushing, 1971; Linsky, 1970; Gove and Howell, 1974; Gove and Fain, 1975). Thus labelling perspectives take the position that what matters most for outcome are differences in social resources or social power; and the variables of class, marital status and work performance are related to outcome because they index resources or power.

An example of the second type of perspective relevant here is the argument that role demands and expectations directly influence rate of recovery or improvement. As Gittelman-Klein and Klein (1968:289) point out, "... [T]he married patient must live up to his family's anticipation that he return to the community as an active citizen, and that his better outcome is a reflection of his response to external support, pressures and expectations."

From the selection artifact point of view such relationships are seen as resulting from differing degrees of disorder or disability. The more symptomatic and/or ineffective an individual, the less likely he will find a marital partner, achieve higher class levels, and work consistently, and the more likely he will spend extended periods in the hospital.

It is important to recognize that this causal association vs. selection artifact interpretive issue is almost entirely without implication for competing theories about the origins of schizophrenia. For example, while the selection interpretation of these relationships with psychiatric outcome is consistent with biological hypotheses, it may be viewed with equal plausibility as reflecting the effects of various social or experiential phenomena. Thus, sociological theories emphasizing the personal consequences of inequities in social systems; social psychological, developmental or cultural theories that postulate the effects of various factors in the socialization process; and environmentally oriented as well as biologically oriented psychiatric theories are all tenable. Indeed, the only position that is clearly not tenable under a selection artifact interpretation is the societal reaction perspective. What sets these diverse positions off from societal reaction theory and makes them consistent with the selection interpretation is their common assumption that effects or consequences, whatever their source, are ultimately expressed in observable differences in social adaptation or social disability, *the locus of which is in the individual*.

Social disability is closely related to the concept of social competence that has recently emerged as a central and useful dimension for conceptualizing mental health (Gladwin, 1967; Phillips, 1968; Smith, 1968; Hersch, 1968; Vance, 1973) and for understanding the occurrence of psychological disorder (Zigler and Phillips, 1960; 1962; Rae-Grant et al., 1966; Phillips, 1966-1967). Competence also has been emphasized in relation to adequacy of role performance (Inkeles, 1966; 1968; Hunt, 1968; Smith, 1968) and as a major factor in prognosis or outcome (see Zigler and Phillips, 1961; Turner and Zabo, 1968). Although there is some variation among conceptualizations of competence, most are associated with developmental theories. In general, these theories assume that individuals progress through a series of stages or levels of maturity and differ widely in the final level attained. The developing individual is confronted

with both problems and tasks posed by society, and deals with them with varying degrees of success. The level of maturity or competence ultimately achieved presumably depends upon how successfully problems have been resolved and tasks accomplished in the course of development. "Social competence . . . corresponds to the individual's level of psychosocial development" and refers to the ". . . ability to meet expectations set by society" (Phillips, 1966:472). Thus competence describes an acquired capacity for socially effective behavior and, as such, resides within the skin of the individual.

Although there is evidence for an important interdependence between competence and type of severity of psychopathology (Zigler and Phillips, 1960; Phillips et al., 1966), the two dimensions are separable both conceptually and in terms of their relationships to outcome (Achenbach, 1966; Cumming, 1963; Garmezy, 1970; Turner, 1977). Thus Myers and Bean (1968:117) argue that "although both role performance and mental status certainly measure social functioning, neither, by itself, seems to be a sufficient measure . . . for the two are not necessarily related at all points." They note: "Epidemiological studies have shown that some individuals are psychiatrically incapacitated, yet they perform their social roles. . . ." Similarly, Neff (1968:98) indicates that there are ". . . numerous instances of frank psychotics . . . with relatively unimpaired work capacities."

The two variables of psychopathology and competence or social effectiveness, taken together, are reasonably (if somewhat narrowly) representative of a traditional psychiatric perspective (see Buss, 1966). However, an emphasis upon one or both is also consistent with the whole array of theoretical perspectives enumerated above. Thus competence and severity of psychopathology may be taken as central variables for representing the selection artifact hypothesis as an explanation for the persistent relationships with which this paper is concerned.

From the foregoing discussion, it is clear that a resolution of the causal association vs. selection artifact interpretive

dilemma cannot support any single specific theory about the course and outcome of disorder to the exclusion of all others. The issue of whether these associations are causal or secondary, however, does have compelling implications for the relative plausibility of *certain* competing theoretical perspectives, including that of societal reaction, and the failure to consider this issue has sometimes damaged the nature of the debate. For example, in a study addressed to the relative importance of psychiatric symptoms and family attitudes as determinants of length of psychiatric hospitalization, Gove and Fain (1975) also consider the variable of marital status. Having concluded that symptoms are causally related to time in hospital while family attitude is probably not a significant determinant, they report that marital status is also strongly and independently associated with length of hospitalization. Gove and Fain (1975:416) see this latter finding as less consistent with the psychiatric perspective than with societal reaction theory. Their apparent uneasiness about this circumstance then leads them into efforts to explain the finding away.² The point to be made here is that the marital status relationship can only be seen as inconsistent with the psychiatric perspective, however narrowly defined, if one ignores the selection artifact hypothesis. It

² A part of their argument is that marital status probably contains much less measurement error than does their measure of psychiatric symptoms, a circumstance that would artificially inflate the marital status-time in hospital relationship relative to that observed between psychiatric symptoms and length of hospitalization. While it must be granted that there is little error involved in measuring marital status, it is not marital status that is the goal of the measurement. Rather, marital status is employed as an operational definition for level of social resources. The notion that there is much more error in measuring symptoms by psychiatrists' ratings than in measuring level of social resources by determining a subject's marital status strikes us as very questionable. Gove and Fain, of course, are not responsible for the tendency to see marital status as an index of social resources. This probably arises from post hoc explanations by societal reaction theorists of epidemiological findings. In our view, the importance of social resources for psychiatric outcome cannot be assessed effectively under the assumption that status on one or more demographic factors provides a valid index of that variable.

is at least as plausible to interpret the marital status-time in hospital relationship as an artifact of the individual's premorbid condition as it is to interpret it from the societal reaction perspective. Indeed some studies have shown that marital status is not significantly associated with outcome when premorbid factors are controlled (Gittelman-Klein and Klein, 1968; Rosen et al., 1971).

While Gove and Fain (1975) fail to consider a selection artifact interpretation of their data, Zigler and Phillips (1961) wholly ignored the social causation alternative in their work. Defining competence as the individual's level of psychosocial development, these researchers devised a competence index that includes the variables of marital status, occupational level, education and work performance. The use of this index assumes that the elements it consists of are symptomatic of a condition within the individual—that there is a process of selection by marriage, by educational attainment, by occupational achievement, and by work performance in which those of low competence are unable to compete and thus are sifted out. Nowhere do these authors entertain the plausible view that the demonstrated association between their index and time spent in hospital may derive from direct social consequences rather than a selection process.

The specific purposes of the present study are: (1) to more fully assess the relationships, including independent associations, between outcome, as indexed by time spent in hospital, and the variables of social class position, marital status and work performance; (2) to evaluate the relative plausibility of the causal association and selection artifact hypotheses for interpreting these relationships; and (3) to evaluate the implications of our findings with respect to specific competing theoretical perspectives.

Our strategy employs multiple regression analyses involving the simultaneous consideration of class, marital status and work performance along with several additional relevant variables. These additional variables include an estimate of social competence and a measure of severity of psychopathology which, we have

argued, represent central explanatory factors within the selection artifact interpretation.

METHODS

The data to be presented here are part of a long-term investigation of the social and psychiatric correlates of successful performance and community tenure among men diagnosed as schizophrenic. The sample was drawn from the Monroe County (New York) Psychiatric Case Register (Gardner et al., 1963). Since its inception in January 1960, the Register has recorded almost all (95%) psychiatric contacts, whether diagnostic or treatment, inpatient or outpatient, public or private, that occurred within the county. Using the Psychiatric Case Register as a sampling base, a random sample of white, male schizophrenics, ages 20–50, was drawn, of whom 212 were later interviewed. A total of 82 cases (27.7% of the total sample) that met all sample criteria were not interviewed.³ Most of these were patient or family member refusals, while a small proportion of the cases could not be traced.

The sample was limited to those patients who had been reported to the register for the first time between January 1, 1960 and June 30, 1963, had no history of psychiatric hospitalization prior to the initially reported contact, were living within the five-county area surrounding Monroe County, and had received a diagnosis of schizophrenia on one or more psychiatric contacts during the 3½ year period. Limited only by these criteria, we thus began with a representative sample of all men who were, on any occasion, diagnosed as schizophrenic. The total psychiatric experience of included subjects varied from a single outpatient visit to many lengthy periods of hospitalization.

Data

The data sources for this study included the Psychiatric Case Register, two sepa-

³ A detailed analysis of these lost cases suggested that this sample mortality does not attach appreciable doubt to the findings to be reported (Turner et al., 1970).

rate and independent ratings by psychiatrists, a lengthy questionnaire administered to subjects by a trained interviewer, and a social worker's report of an interview with a family member. Information taken from the Case Register included all periods of inpatient care. This provided a three to six year record of treatment experience that spanned two to four years prior to interview and one to two years post interview.

The questionnaire covered such areas as social class and mobility, education, utilization of leisure time, living situation, and also gathered detailed information on each subject's work history, work performance, and current work situation.

Before turning to our results, a brief discussion of the more central variables involved in our analyses is in order.

Outcome or Prognosis

Consistent with the majority of previous research, outcome is indexed in relation to time spent in hospital. Accordingly, only that subset of patients who had experienced a significant hospitalization, defined as fifteen or more consecutive days, are considered. Our measure of time in hospital is based upon the total number of days of inpatient care from first hospitalization to the end of the follow-up period three to six years later. Since the length of time between first hospitalization and the cutoff point varied from patient to patient, it was necessary to make an adjustment for such variable risk periods. The time spent in hospital measure was thus calculated as the total number of days in hospital divided by the number of days at risk. Because the distribution of these scores showed a substantial positive skew, a logarithmic transformation of time in hospital values was employed to meet more adequately the underlying assumptions of correlational and regression procedures.

Marital Status

Information on marital status obtained in the questionnaire was confirmed by social workers in their interview of a family member. For analytic purposes the

categories of married, single, and previously married are used, with the latter category including the divorced, the separated and one widowed subject. To treat marital statuses like the nominal categories they are, dummy variable techniques are employed wherever the variable of marital status is included in our analyses.

Social Class

Our estimate of social class position is based upon the respondent's current or last held job (if unemployed at the time of interview). The classification of these jobs was facilitated by the gathering of detailed information on occupational title, work situation, and job characteristics. The index employed was derived by coding these data in accordance with Hollingshead's (1957) seven occupational prestige levels. This procedure assigns the value of one to the highest occupational category and seven to the lowest category. The problem of restriction in range should not be a factor with this important variable given the following distribution by prestige level: 1-2 = 6.2%; 3 = 8.5%; 4 = 22.7%; 5 = 21.3%; 6 = 26.1%; 7 = 15.2%.

Work Performance

In this context, work performance refers to persistence rather than quality of performance. In effect, the issue is the subject's capacity to hold a job. Based upon detailed work histories, performance was coded using a slightly modified version of the Occupational-Adjustment Index employed by Myers and Bean (1968:242). This version is a nine-point index running from "continuously employed" (1), to "virtually never employed" (9). Following the Myers and Bean rationale, the intervening steps combine the factors of current employment status, percentage of time employed in past year (corrected for time in hospital), and percentage of time employed in past five years (corrected for time in hospital).

Severity of Psychopathology

After interviewing the patient, each psychiatrist completed a detailed symptom checklist, rated the patient on eight mental status dimensions and assigned an overall severity rating (Hetzeneker et al., 1966). This overall rating was made on a twelve-point scale ranging from minimal impairment (1-3) to marked impairment (10-12). To assess adequately the relationship between various social variables and outcome, it was clearly necessary to extract any joint variability deriving from differences in severity of disorder. For this, and other purposes, we desired an estimate of severity of pathology that was not materially influenced by knowledge of the patient's social functioning or history of hospitalization. Since the psychiatrists conducting the interviews might learn of and be influenced by such facts, a blind rater was employed. This rater, a psychiatrist with substantial clinical experience, scored the twelve-point pathology scale for each patient on the basis of the symptom checklist completed by the interviewing psychiatrist. Aside from the symptom profile, the only information known to the blind rater was that the patient was a male between the ages of 20 and 50 and that he had on some occasion been diagnosed schizophrenic.⁴ The use of a blind rater provides a measure of severity of pathology that is relatively uninfluenced by knowledge of the patient's level of role performance or history of hospitalization. In the present paper these blind ratings are used in all analyses involving the variable of severity of pathology.

Estimating Social Competence

The problem of how to measure competence effectively is difficult and far from resolved. Although competence and various social achievements are no doubt associated, they are not equivalent, and

efforts to index competence in terms of such achievements cannot be wholly satisfactory (Turner and Zabo, 1968). Competence includes the capacity to perform in socially effective ways, but the attainment of valued statuses depends upon more than just capacity to perform. Among the additional factors that may contribute to status occupancy are differences in access, opportunity, and luck. For example, the amount of competence to be inferred from a given achieved occupational level must depend upon direction and distance moved from point of origin.

Our effort to index competence involves the assumption that variations in intergenerational occupational mobility reflect differences in competence, at least in part. The strategy is to emphasize the competence component through an effort to partial out other factors. A first step is provided by Tumin and Feldman's (1957) Generational Occupational Mobility Score (GOMS) applied to data coded in Hollingshead's seven occupational categories. This score measures mobility in terms of relative achievement. As with most estimates of social mobility the father's position is the base line, but movement is indexed in relative rather than absolute terms. A GOMS score is the *z score* of a subject's occupational level calculated on the basis of the mean and standard deviation of the occupational scores of all subjects and subjects' brothers whose fathers were of the same occupational rank. Since achievement is relative to that of others who began at the same point, it is possible for sons of highest-rated fathers to be scored as upwardly mobile and for sons of lowest-rated fathers to be shown as downwardly mobile. Moreover, because of its relative nature, the measure effectively controls for differential ease of movement at different status levels. The GOMS score thus thoroughly controls the effects of origin status on occupational achievement.

One of the most important predictors of occupational level is education (Blau and Duncan, 1967; Jencks et al., 1972). Although educational attainment must involve some element of competence, it also reflects important differences in such factors as access, opportunity and luck.

⁴ To check on the reliability of these ratings, a second blind rater was employed on the first 100 cases. Using Robinson's (1957) "A," our blind rater agreed with the second blind rater 0.87 and with the psychiatrist whose checklist was used 0.91. By the same measure, the two interviewing psychiatrists agreed 0.85. Robinson's "A" is related to the interclass correlation by the expression: $r = 2A - 1$.

Moreover, it has been shown that the association between education and occupation derives in major part from formal and informal credentialling processes (Berg, 1970; Jencks et al., 1972). We take the position that competence is most clearly and heavily involved in the variability in social mobility that is independent of educational attainment. If this argument is accepted, variability in social movement that is both independent of origin status and unassociated with education may provide an estimate of competence.

To derive an index emphasizing the competence component in social mobility, a regression equation representing the least-squares line of best fit was utilized to predict Generational Occupational Mobility Scores (GOMS) based on level of education. Actual GOMSs were then subtracted from the predicted scores. The result is a measure of social mobility, expressing degrees of over- and under-achievement, that is linearly unrelated to education and independent of social origin. These residual scores are taken as estimates of competence throughout this paper.⁵

The variability shared with education was taken out linearly both because it made sense to us to do so and because the relationship in these data is, in fact, linear. It should be noted that this process aimed at emphasizing the competence component in mobility involved the removal of only systematic variation. Our competence measure thus includes all of the random and measurement error in mobility. Because of this, observed associations with competence should be interpreted as conservative estimates. We recognize that these residual scores still reflect more than competence and error—that elements of luck, opportunity, etc., likely remain. Our argument is that competence

is importantly involved within these scores and that they therefore provide a reasonable and conservative estimate of that dimension.

What is argued above (and we recognize that the argument is bound to be more persuasive to some than to others) is that the procedures we employed provide a rough estimate of level of social competence. No more and no less is claimed. The intention of this procedure was to identify all of the variance in attained status associated with origin and education as one part of the variability in social class level. This comprises 47% of the observed variation. The other 53% of the variation in R's social class position is orthogonal to origins and education and remains unexplained by the stratification process. This represents status achieved independent of the stratification process. People often achieve more or less than would be predicted on the basis of their origins or education. One of the factors that has been judged to be largely influential in explaining this part of occupational attainment is individual competence, a factor that is also thought by many to be an influential factor in schizophrenic outcome.

Looked at in another way, some of the variation in achieved status among men of the same origin status arises from differences in education (largely through the significance of associated credentials) and some is associated with differences in luck, competence, etc. When that variance shared with education is removed we are left with variance in which competence is most clearly and heavily involved. In discussing the unexplained portion of the variance in occupational status, Jencks et al. (1972:131) state,

yet we find it hard to believe that all of this variation is due to luck or chance. Experience suggests that there are personality differences between people who end up in high and low-status occupation. . . . We believe, though we cannot prove, that these noncognitive traits explain part of the variation in adult success.

Two points: (1) this view is quite consistent with our competence interpretation and (2) our results will provide evidence that this variability is not wholly, or even

⁵ While this use of a residual score may seem extraordinary, it is practically and mathematically identical to the more commonly employed procedure of semipartial or part correlation. Whenever the residual is employed in an analysis it is as though a part correlation controlling the effects of education were simultaneously accomplished. Semipartial or part correlation, of course, is a slightly more conservative control procedure than is partial correlation.

Table 1. Intercorrelations of Independent Variables and Time Spent in Hospital

Variable	1	2	3
Single	...		
Work Performance	.372	...	
Social Class Level	.266	.352	...
Hospitalization (log)	.296	.392	.338

largely, chance or luck since random variation cannot, by definition, be *reliably* associated with anything, including time spent in hospital.

RESULTS

Our initial analyses were addressed to the question of whether the persistently reported relationships between marital status, social class, work performance and time in hospital could be observed within the present sample. The correlation matrix in Table 1 shows that the three social factors are correlated with our outcome measure and that their degree of association is roughly similar. The relationships reported from the array of studies reviewed earlier are thus also observed in our data.

With these well-known associations confirmed, we addressed a question that, to our knowledge, has never been fully assessed. There appears to be no evidence from previous research on whether any of the three social factors are associated with outcome independent of the other two. Multiple regression analysis (Table 2) showed significant independent associations for social class and work performance but not for marital status when the inclusion level is set at $p \leq .05$. The directions of these relationships are as expected with better work performance and higher class levels being associated with less time in hospital.

The relationship between marital status and time in hospital observed in this sample appears to derive from variance shared with the other two persistently related variables. This conclusion, however, is highly tentative. While it is unlikely that more than two of these social factors are independently associated with time in hospital, the observation that it is the marital status relationship that is artifactual

Table 2. Determinants of Time Spent in Hospital: Multiple Regression Analysis

	Variables in the Equation			
	r*	Beta	Standard Error of Beta	
Work Performance	.392	.312	.046	
Social Class Level	.338	.228	.092	
R=.446 R ² =.199 F=14.668 p<.0001 n=121				
	Variables Not in the Equation			
	r*	Partial**	F	Sig.
Single	.296	.145	2.516	.115

* Zero-order correlations with log. time spent in hospital.

** Partial correlations with log. time in hospital, controlling for variables in the equation.

derives from statistically nonsignificant differences in the respective zero-order correlations of these variables with time spent in hospital (Gordon, 1968). It is therefore possible that other studies that include this set of variables will observe significant independent effects for marital status.

Additional regression analyses, along with the design controls on race and sex, demonstrate that work performance and social class are connected with time in hospital independent of one another, of marital status, of severity of pathology, of age, race and sex, and of class of origin. This confirmation of independent effects, however, does not assist in resolving the central interpretive problem. Both causal association and selection artifact remain viable hypotheses for explaining the observed independent associations.

It was argued earlier that competence represents an appropriate and central variable for assessing the selection artifact alternative. Accordingly, this variable was introduced into our analyses. We began by considering the social class relationship in the context of the following argument: to the extent that it is social power and/or social resources that account for the relationship, class level should be related to hospitalization with competence controlled; to the extent that it is difference in competence that matters, com-

Table 3. Intercorrelations: Measured and Constructed Independent Variables

	Single	Work	Social Class Level	1	2	3
Predicted Class Level	.065	.094	.684			
Competence	-.303	-.394	-.729	.000		
Pathology	.399	.445	.197	.086	-.190	
Age	-.513	-.239	-.068	.003	.096	-.120

petence should be related with class level controlled.

Partial correlations were calculated as a preliminary evaluation of this issue. Class was found to be unrelated to time in hospital with competence controlled ($r = .03$, $p > .70$), while competence was substantially related to time in hospital with class controlled ($r = .24$, $p < .009$). While these results clearly favor a selection interpretation, our strategy calls for the simultaneous consideration of those major variables that have previously been shown to be importantly related to time in hospital.

Table 3 extends the correlation matrix shown in Table 1 by the addition of age,⁶ severity of pathology, competence and predicted class level. These latter two variables effectively divide our measure of class into two orthogonal parts: (1) our measure of competence which contains all of the random error, and probably other

nonsystematic factors such as luck, along with variance that we believe reflects competence; and (2) that part of occupational attainment derived from or in any way associated with origin status or education. This partitioning of social class is employed in the multiple regression analysis presented in Table 4.

It may be seen that, although this analysis considers all variables we have discussed, when the inclusion level is set at $p \leq .05$, only competence and severity of pathology are entered. Moreover, the R^2 (.240) is equivalent to that achieved when severity of pathology is entered with social class, work performance and marital status ($R^2 = .235$). By far the best predictor in this equation is competence, which has a coefficient that is substantially greater than that for social class shown in Table 2 ($\beta = .346$ vs. .228). That this occurs, given that our competence measure is but a particular portion of the variability in our measure of social class, requires explanation. Analyses indicate that education is a powerful suppressor variable that hides the actual ef-

⁶ This variable is actually age at first contact with the psychiatric register. However, since this is a study of a cohort of first contacts, age and age at first contact are virtually equivalent variables.

Table 4. Determinants of Time Spent in Hospital: Social Power or Social Competence

Variable	Variables in the Equation			
	r^*	Beta	Standard Error of Beta	
Competence	-.400	-.346	.081	
Pathology	.353	.287	.081	
$R = .490$	$R^2 = .240$	$F = 18.62$	$p < .001$	$n = 118$
Variable	Variables Not in the Equation			
	r^*	Partial**	F	Sig.
Age	-.041	.031	.114	.736
Predicted Class Level	.066	.048	.272	.603
Work Performance	.392	.175	3.70	.057
Single	.296	.098	1.14	.288

* Zero-order correlations with log. time spent in hospital.

** Partial correlations with log. time spent in hospital, controlling for variables in the equation.

fects of achieved position on time in hospital. Thus, despite the fact that education and time in hospital are independent in these data, controlling on education significantly raises the relationship observed.

It appears that the variation in class that is associated with time in hospital, other factors controlled, is contained in the competence measure. Stated another way, although competence as measured accounts for only about half of the variability in class level, it accounts for *all* of the variability in class level that is associated with time in hospital.

Although work performance might be considered in future research as a possible independent factor, when competence and severity of pathology are entered with these data none of the other factors shown in Tables 3 and 4 are significant and none of the first-order interaction terms contribute significantly to the prediction of time spent in hospital. The result is a greatly simplified yet equally effective equation that is also quite amenable to interpretation. It includes only those two variables that are most clearly identifiable with the selection artifact interpretation.

However, it might be objected that the developed competence measure does not thoroughly control all factors that might reflect differences in social resources or power—that the observed relationship may reflect the effects of income. Although we lack the data to test this hypothesis, there are at least three persuasive arguments against it. First, income and occupational level have shown surprisingly little connection. As Jencks et al. (1972:226) conclude on the basis of very adequate data, "The earnings of men in the same occupations are almost as unequal as the earnings of random individuals." Under this circumstance it is difficult to accept that the association between occupational level and time in hospital could derive from associated differences in income.

Second, the variable involved must not only account for the class relationship but for the other persistent relationships considered, including that with marital status. Not only is the notion that the marital status-outcome relationship derives from

income differences counterintuitive, there is independent evidence that this relationship arises from associated differences in premorbid social adequacy (Gittelman-Klein and Klein, 1968; Rosen et al., 1971). This evidence, of course, is wholly consistent with our social competence interpretation. Finally, the relationships evaluated in this paper have been observed for state mental hospital populations where the variation in available income resources cannot be great and where, in any event, the contention that time in hospital is associated with variations in social power comprised of income differences, is hard to credit. While we cannot prove that income is not operating in these results, we can find no reason whatever to believe that it is.

In our view our results constitute convincing evidence in favor of the selection hypothesis for interpreting the well-known relationships that are the subject of this paper. From these findings then, there is no reason to believe that there is anything about marriage, or higher class status, or working, *per se* that matters importantly for outcome as measured by time in hospital. Rather, these relationships appear to arise as a consequence of conditions within individuals that are describable in terms of variations in psychopathology on the one hand and capacity for socially effective behavior on the other. The more symptomatic and/or ineffective an individual the less likely he will find a marital partner, achieve higher class levels and work consistently and the more likely he will spend extended periods in the hospital.

Since our results support a selection artifact interpretation, they are consistent with biological theories and with the whole array of environmentally oriented perspectives that emphasize the personal consequences of developmental or social experience. Our findings, on the other hand, substantially contradict the societal reaction interpretation along with other causal association hypotheses and are more consistent with a growing literature that is critical of certain societal reaction explanations (see Gove, 1975).

In drawing this latter conclusion we do

not contend that none of the processes pointed to from these perspectives exist or are not important. To the contrary, there seems good evidence to support the view that factors such as the needs, resources and attitudes of significant others may importantly influence psychiatric outcome.⁷ Moreover, we believe that any psychiatrist with inpatient services experience would readily confirm that social contingencies can be important elements in both admission and discharge considerations. What we do contend is that, however important these factors may be, the proposition that they underlie such stable relationships as those between time in hospital and marital status and social class must be rejected.⁸

This conclusion, the importance of competence that is emphasized in these results, and the relative independence of competence and severity of pathology as predictors of outcome, also have both treatment and research implications. These findings encourage the growing tendency to see the enhancement of competence as a critical and distinct therapeutic goal that, from the perspective of desired outcomes, deserves as much attention in the treatment process as do the dynamics underlying symptom expression. They also suggest that a reorganization of the patient's current role relationships and circumstances may not necessarily influence the course of his disorder. Specifically, these results are contrary to the view that getting married may reduce the likelihood of relapse in schizophrenic men.

While this study does not challenge the

practical utility of marital status or social class position as prognostic indicators, our results suggest that the measure employed here to estimate competence would be a more effective indicator than these variables and probably would be more effective than the Elgin Prognostic Scale. It should again be emphasized that the procedure we have employed can only provide a rough estimate of level of social competence. The further assessment of the role and importance of this construct will require the development of more direct measures, as will efforts to validate the measure used here.

Finally, from the perspective of social psychiatry, this research points toward increased efforts to discover those social processes, associated to varying degrees with location in the social system, that play a strategic role in the determination of both disorder and chronicity. There seems reason to argue that such research should focus upon the effort to understand the social contingencies involved in the distribution of competence.

A Sampling Consideration

While this is a study of men diagnosed as schizophrenic, the fact that only a single such diagnosis was required for inclusion in the study, raises the question of whether all included subjects can be regarded as schizophrenic and whether, therefore, the findings reported apply to that category of subject.

To answer this question the analyses we have reported were repeated on a sub-population composed of only those subjects for whom the diagnosis of schizophrenic was independently confirmed by both study psychiatrists on the basis of a clinical interview ($n = 98$). Although the variance accounted for by the final equation was somewhat lower for this refined population than for the total sample (.20 vs. .24), the results are effectively identical. The beta for competence was .356, while that for pathology was .254 (compare with Table 4). The findings of this study, therefore, can be applied with confidence to schizophrenia. Moreover, it is our opinion that they apply to a much

⁷ For an intelligent review and critique of the body of research that illustrates this point, see Kreisman and Joy (1974).

⁸ There is one caveat attached to this general conclusion. When single men are studied separately the correlations of competence and predicted prestige with time in hospital are -.24 and .25, respectively, while for the sample as a whole they were -.40 and .07. While competence is still of importance for this group, these results are also consistent with the view that there is something about class position per se that matters for time spent in hospital, at least for single men. This "something" might include differences in social resources or power, as is sometimes suggested.

wider population of the psychologically disordered.

SUMMARY

This study has focused upon the well-known relationships of marital status, social class position, and work performance to psychiatric outcome as measured by time spent in hospital. These relationships are well known both because of the persistence with which studies have reported them and because of the provocative hypotheses that have been generated to explain them. The goal of this research has been to contribute to the understanding of the nature and meaning of these relationships.

We began by verifying the relationships within this representative sample of men diagnosed as schizophrenic. Going beyond previous research, significant independent associations were found for social class and work performance, but not for marital status. Indeed, it was demonstrated that social class position and work performance are connected with time in hospital independent of one another, of marital status, of severity of pathology, and of education, age, and class of origin.

To address the central and unresolved question of how these relationships are to be interpreted, a measure of social competence was added to the analyses. As a variable that is closely identifiable with a social selection perspective, the competence dimension provides one basis for assessing the relative plausibility of causal association and selection artifact hypotheses. Multiple regression analysis produced a simplified yet effective equation that included only severity of pathology along with competence. No other factors and no interaction terms made any contribution to the prediction of time in hospital once competence and severity of pathology were considered.

From these results we conclude that the most plausible interpretation of the well-known relationships that are the subject of this paper is that they constitute artifacts of a social selection process. More specifically, our findings indicate that marital status, social class, and probably work performance appear to matter for outcome

largely because varying statuses on these factors reflect important differences in social competence—a dimension that is of crucial significance for outcome.

REFERENCES

- Achenbach, T. M.
1966 "The classification of children's psychiatric symptoms: a factor analytic study." *Psychological Monographs* 80:1-37.
- Berg, Ivor
1970 *Education and Jobs: The Great Training Robbery*. New York: Praeger.
- Blau, Peter and Otis Dudley Duncan
1967 *The American Occupational Structure*. New York: Wiley.
- Brooke, E.
1957 "A national study of schizophrenic patients in relation to occupation." Congress Report, Vol. 3. Second International Congress for Psychiatry, Zurich.
- Brown, G. W.
1959 "Experiences of discharged chronic schizophrenic patients in various types of living groups." *Milbank Memorial Fund Quarterly* 37:105-31.
- Brown, G. W., E. M. Monck, G. M. Carstairs, and J. K. Wing
1958 "The posthospital adjustment of chronic mental patients." *Lancet* 2:685-9.
- Buss, A. H.
1966 *Psychopathology*. New York: Wiley.
- Carstairs, G. M., W. L. Tonge, N. O'Connor, and L. E. D. Barber
1955 "Changing population of mental hospitals." *British Journal of Preventive Social Medicine* 9:187-90.
- Chapman, L. J., D. Day, and A. Burstein
1961 "The process-reactive distinction and prognosis in schizophrenia." *Journal of Nervous and Mental Disease* 133:383-91.
- Cohen, Leon
1955 "Vocational planning and mental illness." *Personal Guidance Journal* 34:28-32.
- Cooper, B.
1961 "Social class and prognosis in schizophrenia, I." *British Journal of Preventive Social Medicine* 15:17-30.
- Cumming, John
1963 "The inadequacy syndrome." *Psychiatric Quarterly* 37:723-33.
- Dunham, H. Warren
1971 "Sociocultural studies of schizophrenia." *Archives of General Psychiatry* 24:206-14.
- Farina, A., N. Garnezy, and H. Barry
1963 "Relationship of marital status of incidence and prognosis of schizophrenia." *Journal of Abnormal and Social Psychology* 67:624-30.
- Gardner, E. A., H. C. Miles, H. P. Iker, and J. A. Romano
1963 "Accumulative register of psychiatric services in a community." *American Journal of Public Health* 53:1269-77.

- Garnezy, Norman
1970 "Vulnerable children: implications derived from studies of an internalizing-externalizing symptom dimension." Pp. 212-39 in J. Zubin and A. M. Freedman (eds.), *Psychopathology of Adolescence*. New York: Grune and Stratton.
- Gittelman-Klein, Rachel and Donald F. Klein
1968 "Marital status as a prognostic indicator in schizophrenia." *Journal of Nervous and Mental Disease* 147:289-96.
- Gladwin, T.
1967 "Social competence and clinical practice." *Psychiatry* 30:30-43.
- Gordon, Robert A.
1968 "Issues in multiple regression." *American Journal of Sociology* 73:592-616.
- Gove, Walter R.
1975 "Labelling and mental illness: a critique." Pp. 58-81 in W. Gove (ed.), *The Labelling of Deviance: Evaluating a Perspective*. New York: Sage.
- Gove, Walter R. and Terry Fain
1975 "The length of psychiatric hospitalization." *Social Problems* 22:407-19.
- Gove, Walter R. and Patrick Howell
1974 "Individual resources and mental hospitalization: a comparison and evaluation of the societal reaction and psychiatric perspectives." *American Sociological Review* 39:86-100.
- Hardt, R. and S. Feinhandler
1959 "Social class and mental hospitalization prognosis." *American Sociological Review* 24:812-21.
- Harrow, M., G. Tucker, and E. Bromet
1969 "Short-term prognosis of schizophrenia patients." *Archives of General Psychiatry* 21:195-202.
- Hersch, Charles
1968 "The discontent explosion in mental health." *American Psychologist* 23:497-508.
- Hetznecker, W., E. A. Gardner, C. L. Odoroff, and R. J. Turner
1966 "Field survey methods in psychiatry: a symptom checklist, mental status scale and clinical status scale for evaluation of psychiatric impairment." *Archives of General Psychiatry* 15:427-38.
- Hollingshead, August B.
1957 "The two-factor index of social position." Mimeo. New Haven, Ct.
- Hollingshead, A. B. and F. C. Redlich
1958 *Social Class and Mental Illness: A Community Study*. New York: Wiley.
- Hunt, J. McV.
1968 "Toward the prevention of incompetence." Chap. 3 in J. W. Carter (ed.), *Research Contributions to Community Mental Health*. New York: Behavioral Publications.
- Inkeles, Alex
1966 "Social structure and the socialization of competence." *Harvard Educational Review* 38:265-83.
- 1968 "Society, social structure, and child socialization." Pp. 73-129 in J. A. Clausen (ed.), *Socialization and Society*. Boston: Little, Brown.
- Jencks, C., M. Smith, H. Acland, M. J. Bane, D. Cohen, H. Gintis, B. Heyns, and S. Michelson
1972 *Inequality: A Reassessment of the Effects of Family and Schooling in America*. New York: Basic Books.
- Kreisman, D. E. and V. D. Joy
1974 "Family response to the mental illness of a relative: a review of the literature." *Schizophrenia Bulletin* 10:34-57.
- Kris, Else B.
1963 "Five-year community follow-up of patients discharged from a mental hospital." *Current Therapeutic Research* 5:451-62.
- Linsky, Arnold S.
1970 "Who shall be excluded: the influence of personal attributes in community reaction to the mentally ill." *Social Psychiatry* 5:166-71.
- Maisel, Robert
1967 "The ex-mental patient and rehospitalization: some research findings." *Social Problems* 15:18-24.
- Mason, A. S., E. K. Tarpy, S. J. Sherman and D. P. Haefner
1960 "Discharges from a mental hospital in relation to social class and other variables." *Archives of General Psychiatry* 2:1-6.
- Myers, J. K. and L. L. Bean
1968 *A Decade Later: A Follow-up of Social Class and Mental Illness*. New York: Wiley.
- Neff, W. S.
1968 *Work and Human Behavior*. New York: Atherton.
- Norris, V.
1956 "A statistical study of the influence marriage has on the hospital care of the mentally sick." *Journal of Mental Science* 102:467-86.
- Phillips, Leslie
1966 "Social competence, the process-reactive distinction, and the nature of mental disorder." Pp. 47-81 in Joseph Zubin and Paul Hoch (eds.), *Psychopathology of Schizophrenia*. New York: Grune and Stratton.
- 1967 "The competence criterion for mental health programs." *Community Mental Health Journal* 3:73-6.
- 1968 *Human Adaptation and Its Failures*. New York: Academic Press.
- Phillips, Leslie, Inge K. Broverman, and Edward Zigler
1966 "Social competence and psychiatric diagnosis." *Journal of Abnormal and Social Psychology* 71:209-14.
- Rae-Grant, Q.A.F., T. Gladwin, and E. M. Bower
1966 "Mental health, social competence, and the war on poverty." *American Journal of Orthopsychiatry* 36:652-64.
- Robinson, W. S.
1957 "The statistical measurement of agree-

- ment." *American Sociological Review* 22:17-25.
- Rosen, B., D. F. Klein, and R. Gittelman-Klein
1971 "The prediction of rehospitalization: the relationship between age of first psychiatric treatment contact, marital status and pre-morbid asocial adjustment." *Journal of Nervous and Mental Disease* 152:17-22.
- Rushing, William A.
1971 "Individual resources, societal reaction, and hospital commitment." *American Journal of Sociology* 7:511-26.
- Scheff, T. J.
1963 "The role of the mentally ill and the dynamics of mental disorder: a research framework." *Sociometry* 26:436-53.
1966 *Being Mentally Ill: A Sociological Theory*. Chicago: Aldine.
- Sherman, L. J., E. C. Moseley, R. Ging, and L. K. Bookbinder
1964 "Prognosis in schizophrenia." *Archives of General Psychiatry* 10:123-30.
- Smith, M. B.
1968 "Competence and socialization." Pp. 270-320 in John A. Clausen (ed.), *Socialization and Society*. Boston: Little, Brown.
- Tumin, Melvin and Arnold Feldman
1957 "Theory and measurement of occupational mobility." *American Sociological Review* 22:281-92.
- Turner, R. Jay
1977 "Jobs and schizophrenia." *Social Policy* 8:32-41.
- Turner, R. J., L. Dopkeen, and G. Labreche
1970 "Marital status and schizophrenia: a study of incidence and outcome." *Journal of Abnormal Psychology* 76:110-6.
- Turner, R. Jay and Lawrence Zabo
1968 "Social competence and schizophrenic outcome: an investigation and critique." *Journal of Health and Social Behavior* 9:41-51.
- Vance, Elizabeth Taylor
1973 "Social disability." *American Psychologist* 28:498-511.
- Wing, J. K., J. Denham, and A. B. Monroe
1959 "Duration of stay in hospital of patients suffering from schizophrenia." *British Journal of Preventive and Social Medicine* 9:145-8.
- Zigler, Edward and Leslie Phillips
1960 "Social effectiveness and symptomatic behaviors." *Journal of Abnormal and Social Psychology* 51:231-8.
1961 "Social competence and outcome in psychiatric disorder." *Journal of Abnormal and Social Psychology* 63:264-71.
1962 "Social competence and the process-reactive distinction in psychopathology." *Journal of Abnormal and Social Psychology* 65:215-22.

THE SOCIOGENESIS OF PSYCHOLOGICAL DISORDER: REEXAMINING THE CAUSAL ISSUES WITH LONGITUDINAL DATA*

BLAIR WHEATON

Yale University

American Sociological Review 1978, Vol. 43 (June):383-403

Although there is some consistency to the finding that socioeconomic status is related inversely to various measures and concepts of psychological disorder, the causal interpretation of this relationship remains uncertain. As usually stated, the issue is whether this inverse relationship arises from the fact and that low social status leads to disorder symptomatology (referred to as social causation) or a high degree of symptomatology impairs the individual's ability to be upwardly mobile (social selection). Research on this issue is far from definitive, especially for less severe types of symptomatology. The causal question is reformulated here in terms of the possible outcomes of a causal analysis of panel data. Based on data from two studies which each measure SES and psychological disorder symptomatology at multiple points in time, results from estimating a series of panel models incorporating unobserved variable specifications of both SES and disorder favor a social causation interpretation. However, differences in results across samples suggest that some modification of the simple social causation position may be necessary.

One of the more consistent and better documented findings in sociology has been the inverse relationship between socioeconomic status and psychological disorder. Dohrenwend and Dohrenwend (1969) report that twenty out of twenty-five studies which use both treated and untreated populations support this finding. Including studies in this list which use only treated samples, or which measure symptomatology in the community, or which have been done more recently, merely adds to the picture of consistency (see Clark, 1948; Hollingshead and Red-

lich, 1958; Sewell and Haller, 1959; Gurin et al., 1960; Haller and Thomas, 1962; Goldberg and Morrison, 1963; Dunham, 1965; Parker and Kleiner, 1966; Turner and Wagenfeld, 1967; Myers and Bean, 1968; Rushing, 1969; Shader et al., 1971; Eaton, 1974; and Wilkinson, 1975). The causal interpretation of this relationship remains uncertain. The question is whether the covariation between SES and disorder variables should be interpreted as arising from a social causation process, a social selection process, or a combination of both. Dohrenwend et al. (1970:197) state the issue this way:

Is low social status more a cause or is it a consequence of psychiatric disorder? On the basis of research to date, it has been impossible to tell: for this relationship can be explained with equal plausibility as evidence of social causation, with the environmental pressures associated with low social status causing psychopathology; or by contrast . . . as evidence of social selection, with pre-existing psychiatric disorder leading to low social status.

Although some evidence has accumulated which suggests that social selection is at least important in accounting for the covariation between SES and schizophrenia (Goldberg and Morrison, 1963; Turner and

* This research was supported in part by the College of Life and Agricultural Sciences, University of Wisconsin, Madison, and the National Institute of Mental Health, PHS Research Grant MH-19689 (Gene F. Summers and John P. Clark, coprincipal investigators); and in part by the United States Office of Education, by National Science Foundation Grant GS29031 (Archibald O. Haller, principal investigator); and by the Manpower Administration, U.S. Department of Labor, Grant 91-55-72-49 (Luther B. Otto, principal investigator). I am indebted to Gene F. Summers for the provision of his data, and I would like to thank Luther B. Otto, Archibald O. Haller, and David L. Featherman for their permission to use the Lenawee County, Michigan data. Gene F. Summers, William H. Sewell, Shalom Schwartz, David Mechanic, David L. Featherman, Jerome K. Myers, and Melvin L. Kohn commented on earlier versions of this paper.

Wagenfeld, 1967), Bruce P. Dohrenwend (1975:386-7) has more recently concluded that the evidence supporting social causation in the genesis of any form of psychological disorder is still slim.

Although the cause-or-consequence issue has changed in form since the pioneering work of Faris and Dunham (1939), the essential causal question involved has survived thirty-five years of research in sociology. Reviews of research in this area mirror the complexity of the causal question (Turner and Wagenfeld, 1967; Robins, 1969; Dunham, 1969) and the intractability of the answer (Mishler and Scotch, 1963; Kohn, 1968; Dohrenwend and Dohrenwend, 1969; B. P. Dohrenwend, 1975). The social selection argument predicts that psychological disorder symptomatology will inhibit upward mobility potential, resulting in lower socioeconomic status over time compared to individuals low in symptomatology; but it usually also involves the presumption that psychological disorder is primarily genetically caused. On the other hand, the social causation alternative presumes that some set of environmental conditions which covaries with SES also determines vulnerability to psychological disorder.¹ This need not be an either-or question, but if the answer clearly favors one position over the other, the social scientist, depending on this answer, will come to very different conclusions about the importance of sociogenic theories of mental health and the viability of social policy in this area.

The social causation vs. social selection argument is more resolvable than the present state of the evidence would suggest. A number of factors in previous research seem to account for the lack of

progress: (1) the problematic, either-or form of the causal question or, alternatively, research designs which allow only an either-or answer; (2) the greater availability of cross-sectional data as opposed to longitudinal data; (3) a tendency to define the causal issue in global terms while only measuring a limited range of psychiatric symptomatology; (4) changes in the definition and measurement of SES over time (5) a reliance on presence/absence notations of disorder rather than more dimensional concepts; and (6) unreliability in the symptom measures.

In this paper data from two different longitudinal studies are used to investigate the plausibility of the social causation and social selection interpretations of the inverse relationship between SES and psychological disorder. The term "psychological disorder" is used specifically to refer to the extent of primarily anxiety- and depression-related psychoneurotic symptomatology manifested by an individual. The extent of this symptomatology indicates the likelihood of a cognitively and affectively based impairment in role functioning capacities.² The proper interpretation of this disorder concept necessarily involves two central features. First, psychological disorder is *not* intended as a blanket term denoting the full range of various types of symptomatology. The data in this study do not allow inferences about the more severe psychotic disorders such as schizophrenia. At the same time, the symptoms measured are used in the definition of many nonpsychotic disorders. Accordingly, the definition above emphasizes this range of symptomatology. Second, disorder is defined dimensionally; no specific symptom level will be used as a criterion to distinguish between illness and its absence. The extent of symptomatology can be considered however, a proxy measure of the changing probability of having sufficient anxiety and depression impairment to qualify for a diagnosis of a disorder of this type.³

¹ There are, of course, a number of alternate versions of both the social selection and social causation hypotheses. Social selection, for instance, is indicated by either actual downward mobility or inhibited rates of upward mobility (Turner and Wagenfeld, 1967). Social causation can be represented via a number of intervening mechanisms: social stress, isolation, socialized rigidity, and perceived goal/means disjunctions are a few examples. However, the different sub-versions of social causation and social selection will not be investigated in this paper. Our interest is primarily in the total effects of SES and disorder variables. Any explication of these effects must await a later paper.

² The term "psychological disorder" is used here, but only as a modal synonym for a number of possible alternative terms (Seiler, 1973).

³ The procedures used to investigate the causal issues in this study could be applied to specific dis-

The whole social causation vs. social selection issue needs to be set in an explicit causal framework. This issue is assessed here through the interpretation of the lagged effects of SES and psychological disorder on each other in panel models which incorporate unobserved constructs and multiple indicators of these constructs wherever possible (Jöreskog, 1973; Jöreskog and Sörbom, 1977).

The State of the Evidence

Most of the studies which have investigated the social causation vs. social selection issue have been concerned with schizophrenia, but these studies are instructive because they illustrate the interpretative issues which arise in this type of research. Evidence for the social causation hypothesis often includes the notion that a sizeable portion of the impaired should come from families of low social status background and/or show a high amount of intergenerational status stability. For example, Hollingshead and Redlich (1958) interpreted the fact that 91% of the schizophrenics in their study were in the same social class as their parents as indirect evidence of social causation. Srole et al. (1962) use a negative relationship between parental SES and an adult psychiatric impairment rating for the same purpose. However, there are alternative interpretations available for the evidence in both of these studies. If social selection was operative in the previous generation in both cases, so that the more impaired parents were overrepresented in the lower class, the effects observed could be due to social selection once removed. Turner and

Wagenfeld (1967) also have observed that the inclusion of education in the social class index in the Hollingshead and Redlich study could be obscuring social selection evidence, since education is a stable component of social class during adulthood.

In contrast with this evidence is a series of studies which argue social selection from data showing that schizophrenics display higher rates of intergenerational downward or inhibited upward status mobility compared to the general population (e.g., Lystad, 1957; Goldberg and Morrison, 1963; Turner and Wagenfeld, 1967). The Turner and Wagenfeld (1967) study is notable because it has been cited as good evidence for social causation as well (Kohn, 1968). Using data from a 1960 psychiatric case register in Monroe County, New York, Turner and Wagenfeld do find that the fathers of schizophrenics are somewhat underrepresented at higher SES levels and overrepresented at lower SES levels, which is consistent with a social causation argument.⁴

Some studies have produced equivocal results supporting neither social causation nor social selection (Dohrenwend and Dohrenwend, 1969; B. P. Dohrenwend, 1975). The Dohrenwends have utilized a novel design that compares advantaged ethnic groups (e.g., Jews, Italians) to disadvantaged ethnic groups (e.g., blacks, Puerto Ricans), and assumes that if social selection is operative we would expect greater numbers of healthy members of disadvantaged groups to remain at lower SES levels and thus lower within-social class rates of disorder among blacks and

orders (including schizophrenia) if the disorder variable was defined in dimensional terms. This dimension would include a known universe of symptoms used to identify the specific disorder, but these symptoms would not be interpreted simply as indicating the presence/absence of that disorder. Rather, one could allow for a range from nonimpairment through preclinical impairment to the clinical threshold for a diagnostically specifiable disorder and beyond to the most severe manifestations of this disorder. This approach could include the notion of a decision tree for deciding between disorders. As each step towards a certain end-point diagnosis is passed, the probability that an individual will receive that diagnosis changes, and the disorder thus could be measured by successive approximations.

⁴ Turner and Wagenfeld in fact conclude that social selection accounts for the major portion of the overrepresentation of schizophrenics at lower SES levels. The emphasis on the social selection interpretation of their data needs to be qualified. Seven occupational prestige levels were used to measure SES—from (1) professional to (7) unskilled. The potential importance of social causation is obscured by the fact that only the extreme level 1 and level 7 categories are used in pointing out that the father's SES distribution is more like the general population's than the offspring patient's distribution. If levels 5 (skilled manual) through 7 are collapsed, it turns out that the fathers are more overrepresented than the schizophrenics. Although this may be overgrouping the data, focusing on the lowest prestige group only may be misleading.

Puerto Ricans. Evidence for social causation would be higher rates of disorder for these groups compared to advantaged groups. Unfortunately, results tend to show that Puerto Ricans have higher rates at each class level but that blacks do not differ from Jews and Italians. One comment is in order regarding this design: it in fact assumes that the issue is social stress vs. social selection. Thus, the causal issue in terms of SES is not addressed directly, and if this design has relevance to the social causation vs. selection issue in terms of SES, it requires the rather uncertain assumption that stress empirically explains the impact of SES on disorder.⁵

The overall picture from the sociological studies reviewed is mixed: some studies suggest social causation, some suggest social selection, some both, and some do not provide evidence for either. There are enough problems with this body of evidence to suggest that the cause-or-consequence issue is far from resolved and a reassessment of this issue is in order.

A Clarification of Issues

The causal issue involving social causation and social selection cannot legitimately be addressed in either-or terms. The important information is the degree to which each process is operative. Both the research design and the method used to analyze the data must allow for the sepa-

rate estimation of social causation and social selection effects and the strengths of these effects. This can be accomplished by using data with measures of both SES and disorder at multiple time-points and by estimating the magnitude of the effect of each variable on the other at a later point in time.

The traditional social causation argument emphasizes parental SES as the causal source variable. Kohn (1968) correctly points out that this is only one form of social causation, since adult status could have an independent causal effect. In this study the effects of father's SES and the respondent's own adult SES will both be considered, but more emphasis will be given to adult status as a causal source. There is some potential for confounding between genetic endowment and either of these status variables. But in interpreting the effects of adult status, background status can be controlled; and at least the unmeasured effects of genetic endowment which are confounded with background status can be accounted for. Although adult status effects may seem somewhat safer to interpret in general, this does not imply that the effect of father's SES is uninterpretable as evidence for social causation. Kohn (1972) has concluded that the potential for confounding between genetic and social class background is not sufficient to explain class differences in the incidence of schizophrenia. In view of the moderate type of anxiety/depression symptomatology measured in this study, and the conclusion from Pollin et al. (1969) and Rosenthal (1970) that the evidence for a genetic basis in the neurotic disorders is less convincing than in the case of schizophrenia, it seems to be all the more unlikely that genetic variables could explain the effects of SES in this study.

The implications of the differences between cross-sectional and longitudinal data have not yet been fully incorporated into etiological research on mental health. In the social selection studies, for instance, the possibility of confounding between social mobility effects and social selection effects has not been addressed (Fried, 1975). The conclusion that inhibited mobility results from impairment is

⁵ This assertion is important to the justification of the longitudinal design used in this study. In the ethnic group mobility design, the groups must vary reliably on the rate of exposure to environmental stressors (including a discriminatory environment) or the results are inherently ambiguous. If this assumption holds, then it must also be true that the amount of stress covaries negatively with SES with sufficient strength to act as the intervening variable which results in a zero net effect of SES on disorder. Stress is only one of a number of possible concepts which could play an explanatory role here. These other concepts (e.g., perceived goal/means discrepancies, social isolation, learned rigidity) should be seen as alternative models, not sub-versions of a stress model. Moreover, when stressful life events (in various forms) is posited as the explanatory variable in a SES-Disorder model, results often show that it does not explain the effects of SES, although at least one study does report that SES has no net effect (Myers et al., 1974).

usually unwarranted; it is possible that the debilitating experience of either downward or inhibited upward mobility is what caused the disorder. This problem is avoided with panel data, since it is possible to measure the existence of disorder before observing any status mobility over time.

Given the mild to moderately severe symptomatology at issue here, use of an untreated community population will be less of a problem than usual, as long as our inferences only include the type of disorder defined above. Community data, however, are particularly subject to validity problems when it becomes necessary to use a cutting point on the measure of disorder to make sick-well distinctions or decisions about impairment (Kohn, 1968). One result of this has been an endless debate over the meaning of different cutting-points in different samples (see the review by Seiler, 1973). As stated earlier, disorder in this study is conceived of as a dimension and it will be retained as such.

In reconciling previous findings on the social causation vs. social selection issue and making predictions for this study, two factors must be considered: (1) the measure of socioeconomic status and (2) the type of disorder concept used. It is consistently the case that studies which include the respondent's education as an indicator of SES emphasize social causation (e.g., Hollingshead and Redlich, 1958; Langner and Michael, 1963), while those which use only occupationally-based measures of status emphasize social selection interpretations of their data (e.g., Goldberg and Morrison, 1963; Turner and Wagenfeld, 1967). Thus, to maximize the chance of detecting a social selection effect, it is advisable to use a measure of occupational prestige only.

It is expected that social causation will play a predominant role in the explanation of the negative covariance between socioeconomic status and psychological disorder in the data used in this study.⁶ What-

ever importance social causation (in terms of SES) has in the case of schizophrenia, social causation effects are likely to be more important relative to selection effects the less severe the type of symptomatology measured. Presumably, it would take more pervasive, more severe, and more constantly visible symptoms than the type of anxiety/depression symptomatology at issue here before any effect on such relatively stable and complex characteristics as SES could be expected. This prediction implies that, as we move from schizophrenia to less severe disorders, the chances for social selection should decrease, but it need not be the case that social causation effects should increase in absolute magnitude as well.

DATA AND METHODS

Results will be compared across three panel samples whose characteristics overlap in varying degrees. The pattern of similarities and differences in the characteristics of these samples will be central to our decision about the cause-or-consequence issue.

Data Set I: Illinois

This study was undertaken in 1966 to monitor the effects of industrial development in a rural area in Illinois as compared to a similar but nondeveloping rural area in another part of the state. The same schedule was used in interviewing a stratified multistage cluster sample of adult heads of household in each of the two regions in 1966. The experimental area—where a high production cold-rolling steel mill was about to be built—will be designated the Hennepin area in the subsequent analyses; the control area

⁶ The assertion that there is a negative relationship between SES and a concept of disorder based on psychoneurotic symptomatology is not trivial. Fried (1975) points out that studies which measure diagnosed or treated neurosis find a positive relationship, while those which use undiagnosed or field interview

measures of neurotic symptomatology find a negative relationship. It is important that a distinction is made between the relationship between SES and the process of labeling symptoms by professionals and the relationship between SES and the existence of such symptoms. The inconsistency noted by Fried (1975) suggests that the fact of entry into a treatment system may alter the apparent nature of the relationship. This could happen, for instance, if there is class-related selectivity into treatment, or if there are therapist biases in diagnostic decisions about severity.

will be referred to as the Watseka area. Follow-ups of the original sample in each area were conducted in 1967 and 1971, and included measures of the anxiety/depression symptoms of interest in this study and the respondent's occupation from 1960 through 1971. The samples will be kept separate because the respective covariance matrices for all SES and disorder measures proved to be significantly different.⁷ The age, sex, education, and income distributions in each sample were extremely similar. The final N in Hennepin was 736, and in Watseka 295.

Data Set II: Michigan

These data are from a longitudinal study of the aspirations and status attainments of a sample of youth from Lenawee County, Michigan. The first wave included all 17-year-old male students in the Lenawee County school system in 1957. Jarrett (1961) points out that the area has an evenly divided farm, rural nonfarm, and urban population. Lenawee County is a more urbanized area on the whole than either Hennepin or Watseka, and the age/sex characteristics of this sample are clearly different as well. A follow-up was conducted in 1972, and is described in some detail in Otto and Featherman (1975). Data were collected by telephone and mailed questionnaire via a concentrated tracking effort. The final N in the panel sample analyzed here is 291.

Measuring Psychological Disorder

The common symptoms measured in the Illinois and Michigan studies are from the Langner 22-item Index (Langner, 1962). This Index has been the source of some controversy (Dohrenwend and Dohrenwend, 1969; Phillips and Clancy, 1970; Seiler, 1973), mainly focusing on validity and response bias problems. The symptoms from this Index are listed in subgroups according to type in the Ap-

pendix (only key phrases in the wording of each symptom are included).⁸ Clearly, the symptoms do not tap a universe which includes psychotic behavior. The problem is to specify a content universe for disorder inductively from inspecting the content of the specific items in the Langner Index. The most visible theme running through these symptoms is an emphasis on anxiety and depression.⁹

Measurement problems with the Langner Index have been extensively investigated, but no study has conclusively demonstrated that these problems obscure relationships between SES and disorder as defined here. Validity studies which compare Langner scores across groups sometimes produce unexpected results—for instance, higher scores among college students compared to predischARGE patients in a hospital (Manis et al., 1963) or among clinic outpatients compared to hospitalized inpatients (Dohrenwend and Crandell, 1970). These studies are subject to a common alternative interpretation. Patients in hospital settings are more likely to be there because of some form of psychotic rather than neurotic behavior. On the other hand, outpatients and specialized groups like college students may be expected to exhibit higher rates of common neurotic symptoms and/or are more amenable to self-report measures of behavior.

The most serious potential response bias problem is one of social desirability content in the items (Dohrenwend and Dohrenwend, 1969). Phillips and Clancy

⁸ The respondent indicates the frequency of occurrence of each of the problems—for instance, in item S the respondent reports how often he feels "somewhat apart" even among friends or in item I, how often he has shortness of breath even without exercise.

⁹ Space does not permit a description of the decision procedure, but it is based on: (1) definition of the symptoms by psychiatrists in the Crandell and Dohrenwend (1967) study; (2) the labeling of similar symptoms in other indices (Schwartz et al., 1973); (3) classification of the symptoms in Coleman (1964); (4) the fact that the questionnaire/interview situation is most suited to the measurement of moderate rather than severe symptomatology and that responses probably indicate moderate impairments which are usually associated with anxiety/depression problems; and (5) correlations over .7 between Langner scores and measures of stable anxiety (Shaffer et al., 1971).

⁷ This test was carried out using the van Thillo and Jöreskog (1970) computer program SIFASP. Constraining the two covariance matrices for Hennepin and Watseka equal produced a significant χ^2 , indicating that the matrices were indeed different and should be analyzed separately.

(1970) investigated the relationship of the social undesirability ratings of the Langner items to SES and Langner symptom scores. Social undesirability scores are related to Langner scores, but the desirability variable only partially accounts for the relationship between status and symptoms. Thus, there is some credence to the assertion that an interpretable relationship between SES and disorder exists, although desirability may attenuate the relationship somewhat.¹⁰ A sensitivity analysis of the effects of nonrandom error in the symptoms will be conducted in the present data.

Only the two psychological and psychophysiological symptom subscales from the Langner Index (see Appendix) can comfortably be used as indicators of psychological disorder. This is due to the fact that, while a sample of psychiatrists and internists agreed that the symptoms on these subscales indicate primarily psychological and rarely organic problems, the physiological subscale was judged to indicate organic problems and there was some disagreement about the ambiguous symptoms (Crandell and Dohrenwend, 1967). The ambiguous symptoms are related to age; thus the problem is more acute in older samples.

Methods

The panel models that will be used to test the cause-or-consequence issue allow for the specification of a causal system among a set of unobserved constructs as well as a measurement specification relating the measured variables to these con-

structs. The specification and estimation issues in this type of model have been discussed at length by Jöreskog (1973; 1976) and by Jöreskog and Sörbom (1977). Optimally efficient estimation of the parameters is achieved via maximum likelihood methods. A χ^2 goodness of fit statistic can be used to assess specification issues, given that the model is overidentified (Jöreskog, 1973).¹¹ Larger values of χ^2 for a model imply poorer fit.

Using an unobserved variable specification of the relationship between SES and disorder results in a number of analytical advantages: (1) certain parameters can be fixed or constrained equal in these models and the consequences on the goodness of fit can be monitored; (2) the true relationship between SES and disorder can be estimated after taking into account random error in the measures; (3) it is possible to conceive of psychological disorder as an unobserved construct with multiple indicators; and (4) the effects of certain types of nonrandom error (e.g., methods effects) can be estimated for multiwave variables which have at least two indicators.

Decisions about Causality in Panel Data

There can be some very real problems with causal inference in panel data (Bohrnstedt, 1969; Duncan, 1969; Heise, 1970; Pelz and Lew, 1970). But it is quite another question as to whether causal inference, defined in a restricted manner, is pragmatically possible and analytically useful in these situations. The particular form of the causal question can simplify the task of inference. Interpretative problems can arise from the direct comparison of the magnitude of cross-lagged coefficients in two-variable panel models. Comparing the unstandardized coefficients makes no sense when the dependent variables are in a different metric, and standardizing does not necessarily help.¹² Accordingly, the question will not

¹⁰ The objection that Langner scores only tap transient and reversible responses to stress has not been addressed in the text, but it should be noted that the stability of particular symptoms over time is less important than the stability of the total number of symptoms. In this regard, three sources of evidence can be cited. First, in the present study, the true score correlation between Langner scores over a four-year period exceeds .6. Second, Haberman (1965) reports a .54 correlation between scores over a three-year period in a New York City sample. Third, Langner Index scores correlate more highly with scales intended to measure stable anxiety than with scales which measure short-term anxiety (Shader et al., 1971). The phenomenon measured by the Langner Index is best described as having moderate stability.

¹¹ Estimation of all models in this study was done with LISREL II (Jöreskog and van Thillo, 1973). The analysis was complete before the appearance of the new LISREL III.

¹² The comparison of standardized coefficients is sometimes meaningful when the variables involved

be asked in terms of whether A causes B more than B causes A, except in the sense defined below.

The inferences in this study are primarily based on simple indications of statistical significance vs. nonsignificance. The outcome categories are that one, both, or neither of the variables involved is causally relevant to the other. The criterion may be arguable, but from the point of view of making an inference the researcher is always faced with the job of deciding whether a coefficient is interpretable. The sense in which we can say that social causation is more important than social selection is in terms of the fact that the probability that the earlier SES in a panel model has some negative effect on later disorder symptomatology is substantially higher than the probability that earlier disorder symptomatology has some negative effect on later SES.

The cross-lagged models used to investigate the cause-or-consequence issue in this study are based on the two-wave two-variable model proposed by Heise (1970), with a number of modifications. Perhaps the most basic assumption in the Heise model is that the concept of causality implies a time lag; therefore as long as the causal lags are properly specified within a certain range (in general the measured lag should not differ from the true lag by more than five lags—see Pelz and Lew, 1970) there is no necessity for the specification of instantaneous reciprocal effects for endogenous variables in the model in addition to the cross-lagged effects. Moreover, when lagged effects and instantaneous effects are not both identified, the finite time lag requirement suggests that the preferred specification of causality is in terms of lagged effects.

Three extensions of the Heise model are incorporated in the present analysis. First, the assumptions that the causal lags of the two variables at issue are equal and that the measurement interval in the data approximates the causal lag can be relaxed somewhat: when there is variance in the

measurement lag within and across data sets as in the present case, there is a corresponding increase in the flexibility of inference procedures. Second, the distorting effects of measurement error can be taken into account by specifying SES and psychological disorder as unobserved constructs with fallible indicators. Third, there must be an attempt to control for confounding variables which may affect the interpretability of the cross-lagged effects.

SPECIFICATION OF THE CAUSE-OR-CONSEQUENCE MODEL IN THE TWO DATA SETS

Model I: Hennepin and Watseka

The basic model for testing the cause-or-consequence issue in the Hennepin and Watseka samples is shown in Figure 1 (paths are lettered to facilitate discussion). Three successive measures of the respondent's SES are included: x_2 through x_4 are Duncan SEI scores in 1960, 1965, and 1971, respectively. A similar measure of background status during adolescence is included as x_1 . Sex and age will be added to the model as perfectly measured exogenous variables. Psychological disorder is measured at two points in time (1967 and 1971), and is indicated by the two psychological and psychophysiological symptom subscales of the Langner Index at each time point (x_5 through x_8); λ_{p1} through λ_{p4} , the factor loadings for each of these subscales, indicate the degree to which the unobserved psychological disorder variable accounts for the scores on each of the indicators.

The crucial paths to compare in Model I are either a vs. b, g vs. b, or possibly j or h vs. b. Each of a, g, j, or h represents a version of a social causation effect (varying the time lag); b stands for the social selection effect. The interpretation of the various outcomes for Model I are shown in Table 1. Each cell of this table stands for one of the four possible outcomes. The high-low distinction for each parameter represents the application of some statistical significance criterion when the parameter is estimated.

Cell 3 in Table 1 is an outcome showing clear social causation and no operative

are measured in an arbitrary metric, but ambiguities in interpretation still may exist if the variances of the two variables of interest are differentially representative of the true variances in the population.

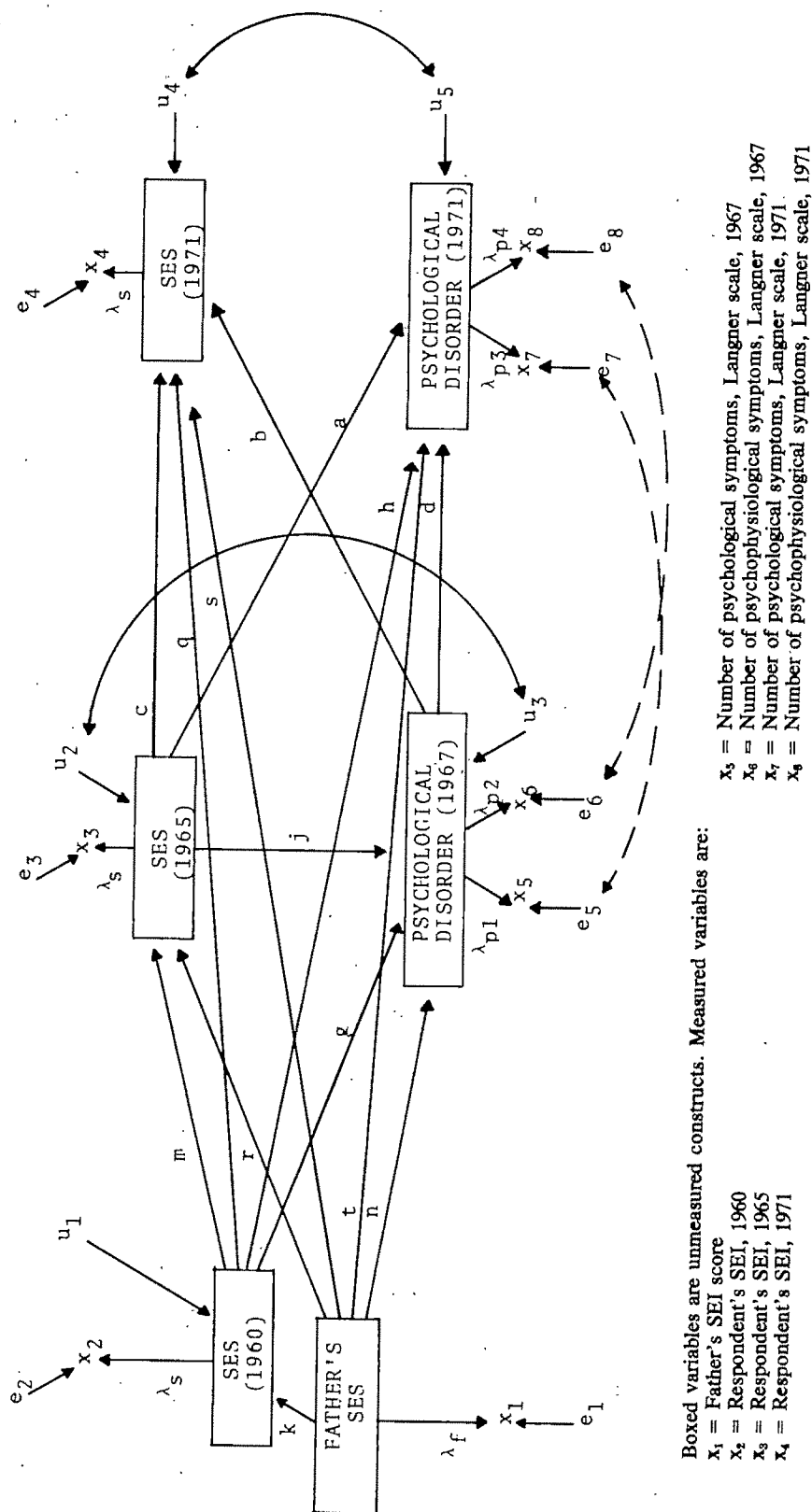


Table 1. Outcomes for the Cause-or-Consequence Issue with Accompanying Interpretations

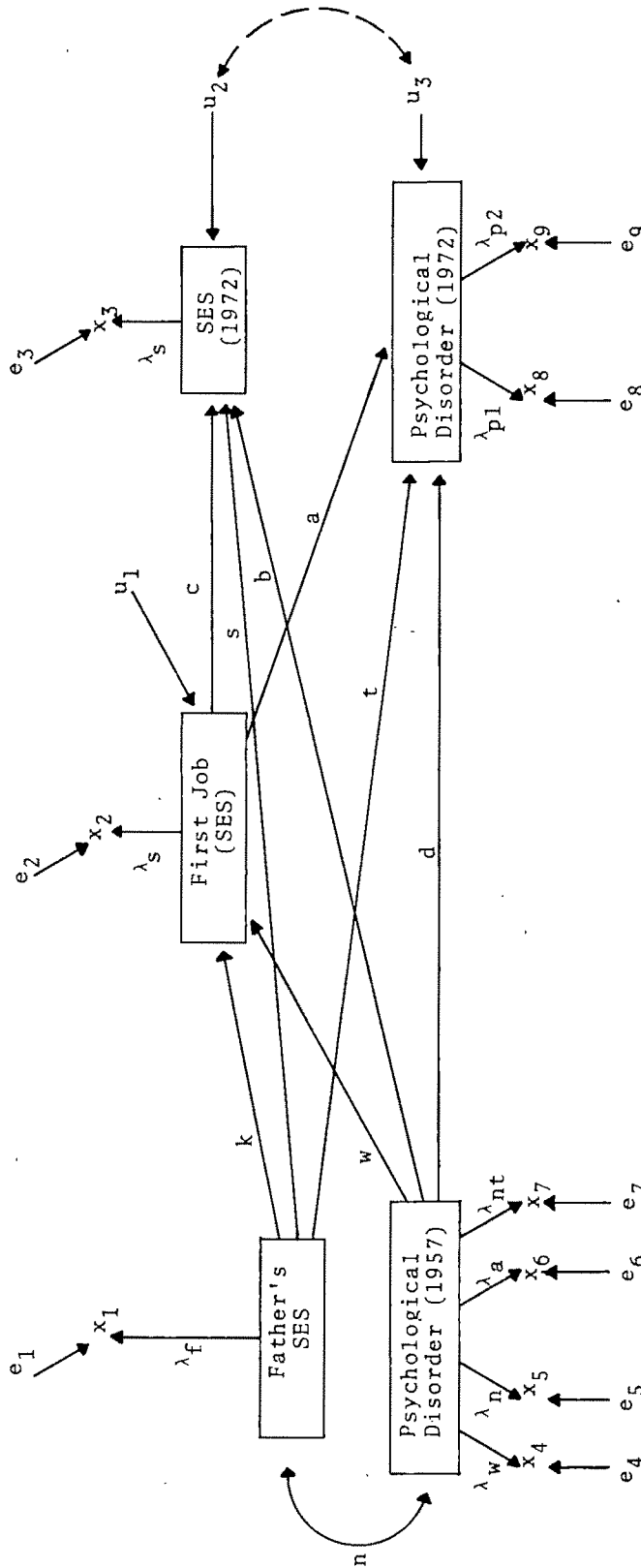
Parameter b (also w in Model II)	Parameter a (also g, j, or h in Model I)	
	High	Low
High	1. social selection+ social causation	2. clearest selection; no social causation
	3. clearest social causa- tion; no selection	4. neither selection nor social causa- tion; possible explanations listed below: (1) other excluded causes con- founding effects (2) high stability of variables (3) methods effects (4) intervening causes excluded (5) nonlinearity (6) specific sample characteristics

selection factors. In this cell, we see that SES in 1965 (or 1960) has a significant negative effect on psychological disorder in 1971 (or 1967), but that disorder in 1967 has no inhibitory effect on later SES. Cell 2 is a clear social selection with no social causation outcome, the opposite finding to Cell 3. Here symptom level in 1967 does affect status in 1971 negatively although status has no effect on later symptoms at any place in the model. Path b is a selection effect because it stands for a retardation in status attainment between 1965 and 1971 attributable to disorder and because the effect is net of previous SES. This leaves an obvious interpretation for Cell 1, i.e., that both social selection and social causation processes contribute to an overall negative relationship between SES and disorder.

This leaves Cell 4—the null cell in terms of the causal issue. A number of possible explanations for a null result are listed in Table 1. Many of these interpretations can be tested. For instance, sex and age can be added to the model to assess further confounding. Variance in the stability of variables across data sets can be used to assess whether social causation or selection effects are apparently being suppressed, nonrandom error correlations (e.g., $r_{es_{67}}$ and $r_{es_{68}}$ in Figure 1) can be estimated, and certain types of nonlinearity can be assessed by entering the indicators in the model in nonlinear form.

Model II: Lenawee County

The comparable causal issue model in the Lenawee County data is shown in Figure 2. Paths analogous to those in Model I for the Illinois data are labelled with the same letters. But there are a number of differences in the specification of Model II which require explanation. This version of the cause-or-consequence model has two instead of three measures of the respondent's SES: first job status (x_2) and the respondent's current SES in 1972 (x_3), both coded in Duncan SEI metric. The psychological disorder measures in 1972 are attenuated versions of the psychological and psychophysiological symptom subscales from the Langner Index (x_8 and x_9 , respectively). Four of the ten psychological symptoms and three of the five psychophysiological symptoms were measured in the 1972 follow-up. Thus, these measures may not exhibit as much reliable variance as the Illinois indices. Two things can be done. First, the psychological subscale is less representative than the psychophysiology subscale, and thus less emphasis should be placed on the psychological symptoms. Second, two ambiguous symptoms (items E and P in the Appendix) can be added to the psychophysiological subscale here because the problems of confounding of the ambiguous symptoms with physical illness are more relevant in older samples (age of



Boxed variables are unmeasured constructs. Measured variables are:

- x_1 = Father's SEI
- x_2 = First job SEI
- x_3 = SEI, 1972
- x_4 = Withdrawing tendencies, California Test of Personality, 1957
- x_5 = Number of nervous symptoms, California Test of Personality, 1957
- x_6 = Anxious insecurity, 16 P. F. Test, 1957
- x_7 = Nervous tension, 16 P.F. Test, 1957
- x_8 = Attenuated psychological symptom subscale, Langner scale, 1972
- x_9 = Attenuated psychophysiological symptom subscale, Langner scale, 1972

Figure 2. Model II: An Assessment of the Cause-or-Consequence Issue in the Lenawee County Data

Lenawee County cohort = 32, 33), and because psychiatrists in the Crandell and Dohrenwend study did indicate that these two items at least were indicative of psychological disorder.

The Langner items were not included in the first wave of the Lenawee County study; thus, it was necessary to use the most nearly equivalent measures in the 1957 data. These measures are: x_4 = withdrawing tendencies, a scale from the California Test of Personality (CTP); x_5 = number of nervous symptoms, also from the CTP; x_6 = anxious insecurity and x_7 = nervous tension, two scales from the 16 P.F. test (Cattell et al., n.d.).¹³

The path a vs. path b comparison in Model II differs from Model I in that the time lags differ. Here the social selection path is over a considerably longer lag (15 years compared to 4 years in Model I). Given this contrast, if $b=0$ in both models, the evidence is stronger that social selection is not operative. Path w in Model II is probably the weakest evidence for social selection in the two models, for at least two reasons. First, there is less certainty about the exact conceptual domain tapped by the 1957 measures of psychological disorder in Lenawee County. Second, this selection effect occurs prior to the actual beginning of the individual's job career, and thus does not bear on the issue of selection during the occupational career. Path w in fact probably reflects a selection effect retarding educational attainment (note that w*c does stand for an indirect selection effect on SES).

RESULTS

Hennepin and Watseka

The correlations and standard deviations for the measures in Model I for both the Hennepin and Watseka samples are

¹³ These measures do share in common with the Langner Index a number of symptoms. The withdrawing tendencies scale is described as measuring feelings of withdrawal and depressive malaise. The nervous symptoms scale contains a number of anxiety and neurasthenic symptoms. The anxious insecurity scale is intended to tap free-floating anxiety, and the nervous tension scale is considered a measure of nervous anxiety and restlessness.

shown in Table 2.¹⁴ Three representations of x_1 are shown; these correspond to re-specifications of the model with different background variables.

Linear models. Because of the collinearity between SES in 1960 and in 1965 (see Table 2), one of the two paths from these variables to each of the other variables in the model was set to zero in all cases ($h=j=q=0$). Other specification conventions are as follows: (1) the factor loadings for the three SEI measures are set equal across time (λ_{a1}), and the error terms are set equal also;¹⁵ (2) father's SES is assumed to be measured without error; and (3) the loadings for the psychological symptom scales (λ_{p1} and λ_{p3}) were set to one (to give the unobserved disorder variable a metric) and the psychophysiological symptom loadings (λ_{p2} and λ_{p4}) were set equal to each other. This specification ensures, in effect, that the same SES and the same disorder constructs are extracted by the model across time.

The results for Model I in Hennepin are shown in Table 3. Father's SES has no net effect on disorder either in 1967 or 1971; thus, there is no evidence for direct intergenerational social causation. If genetic endowment and background status are correlated as a result of some multigenerational social selection process, genetic effects on this type of disorder conceivably could surface via path n or t. The absence of any relationship here suggests that if there are genetic effects on disorder which are confounded with SES, they affect the model in a less obvious manner. Although paths a and b are zero in Hennepin, the social causation path from SES in 1960 to psychological disorder in 1967 is significant and negative ($g^* = -.194$).¹⁶ Given the

¹⁴ The correlations between status and disorder variables in this matrix are quite low, except for some cases in the Hennepin data. It should be remembered it is the correlations between the unobserved status and disorder constructs (as specified in Figure 1) which are of primary interest. There is in general more evidence of negative covariance between the SES and disorder constructs in the Hennepin data than in the Watseka data.

¹⁵ LISREL II could not produce estimates of the model when these error terms were allowed to be free parameters.

¹⁶ The asterisk will be used to denote standardized versions of parameters.

Table 2. Correlations and Standard Deviations for the Cause-or-Consequence Analysis in the Illinois Data *

	X_i										S.D.
	i	ii	iii	X_1	X_2	X_3	X_4	X_5	X_6	X_7	
i Father's SEI119	-.011	.386	.344	.386	.381	-.042	.018	-.037	19.45
X_1 SEX	.020430	.006	.026	.006	-.024	.336	.360	.301	.365
iii AGE	-.130	.240	...	-.040	-.012	-.040	-.112	.114	.222	.231	.310
X_2 SEI (1960)	.257	.059	-.030	.884884	.844	.049	.026	-.007	16.08
X_3 SEI (1965)	.248	.091	-.075882903	.028	.051	-.028	21.32
X_4 SEI (1971)	.264	.124	-.033	.863	.827	.863071	.057	-.049	21.28
X_5 PSYCH (1967) ^b	-.072	.240	.111	-.096	.166	-.096	-.089649	.554	21.89
X_6 PHYS (1967)	-.038	.234	.164	-.016	-.104	-.016	-.030	.454436	1.76
X_7 PSYCH (1971)	-.092	.126	.033	-.088	.120	-.088	-.006	.526	.247653
X_8 PHYS (1971)	-.013	.160	.040	-.011	-.139	-.011	-.005	.377	.309	...	1.70
S.D.	19.98	.380	14.63	22.85	22.82	22.85	22.69	1.45	.555	1.38	.626

* Hennepin sample below the diagonal (N=603), Watseka sample above the diagonal (N=250).

^b In this and in following tables PSYCH refers to the psychological symptom subscale from the Langner Index, and PHYS refers to the psychophysiological symptom subscale from the Langner Index.

reasonably high stabilities of both SES and disorder, the nonsignificant a path can be attributed in part to the g social causation effect in the previous lag. Thus, there is a notable social causation effect in the Hennepin model, but there is no evidence for social selection.

On the other hand, there is no evidence in Table 3 for either social causation or social selection effects in the Watseka sample. One problem with the conclusion that social selection is not operative in these samples is the high stability of SES. What if the model underestimates the social selection causal lag? The Lenawee County data has a longer lag for the selection effect; thus if no selection effect surfaces over the longer lag this will strengthen the conclusion that the null finding here is not simply due to the time lag in the Illinois panel.

The fit of these models can be roughly described by the χ^2 for each model divided by the degrees of freedom—a descriptive fit ratio. Table 3 shows these ratios to be 6.3 and 4.0 in the two samples, respectively, and the probability levels of the models are far beyond .05. These fit ratios indicate the models have a moderate fit: there is certainly room for improvement, but the unexplained residuals produced by the model ($\bar{\Sigma}$, the covariance matrix produced by the model; -S, the input covariance matrix) are still fairly low. Given the null evidence for social selection as well, further specifications should be assessed.

Nonlinear and correlated error models. The possibility that the relationship between SES and disorder may be nonlinear has been suggested by a number of studies (Hollingshead and Redlich, 1958; Srole et al., 1962; Phillips and Clancy, 1970; Shader et al., 1971). Usually the form suggested is best approximated by a negative logarithmic or exponential curve. An inspection of the fit of a number of nonlinear versions of the relationships for paths a and b in Model I suggested that the only modification which might make a difference would be the specification of the relationship between SES in 1965 and psychological disorder in 1971 in Hennepin as:

$$\log Y = b (-1/X).$$

Table 3. Results for the Cause-or-Consequence Issue in the Hennepin, Watseka, and Lenawee County Samples

Parameters: ¹		Hennepin		Watsseka		Lenawee County	
Model I	Model II	U	S	U	S	U	S
(a) SES65-DIS71	FJOB-DIS72	-.013	-.010	-.079	-.061	-.090*	-.120
(g) SES60-DIS67	...	-.240***	-.194	.075	.054
(b) DIS67-SES71	DIS57-SES72	-.001	-.001	.023	.031	.013	.011
(w) ...	DIS57-FJOB	-.227***	-.185
(d) DIS67-DIS71	DIS57-DIS72	.671***	.681	.754***	.813	.252***	.271
(c) SES65-SES71	FJOB-SES72	.921***	.928	.974***	.945	.856***	.906
(m) SES60-SES65958***	.957	.906***	.908
(n) FASES-DIS67	FASES-DIS57	-.002	-.029	-.002	-.030	-.065	-.065
(k) FASES-SES60	FASES-FJOB	5.35 ***	.268	6.88 ***	.354	.442***	.359
(r) FASES-SES65000	.002	.004**	.076
(s) FASES-SES71	FASES-SES72	.002	.035	.001	.016	0. ^f	0.
(t) FASES-DIS71	FASES-DIS72	-.002	-.034	.001	.014	0. ^f	0.
	λ_a	21.90*	.960	20.74*	.973	19.43*	.890
	λ_{p1}	1. ^f	.858	1. ^f	.787
	λ_{p2}	.249 ^b	.547	.389 ^b	.824	1. ^f	1.
	λ_{p3}	1. ^f	.879	1. ^f	.755
	λ_{p4}	.249 ^b	.612	.389 ^b	.799
	$r_{u_{gug}}$.108	.088	.019	.014
	$r_{u_{gu}}$.098	.081	-.066	-.050
	$\chi^2/\text{d.f.}$	69.4/11=6.3		44.3/11=4.0		28/16=1.7	
	Prob. level	.0000		.0000		.0312	

For paths between constructs:

* = $p < .05$

** = coefficients exceeding twice their standard error, $p < .02$

*** = $p < .01$

Codes: U—unstandardized solution
S—standardized solution

f—fixed parameters
a, b—parameters constrained equal

¹ See Figures 1 and 2 for the paths corresponding to the parameters listed. In this table: DIS=psychological disorder; SES=socioeconomic status; FJOB=first job SES; FASES=father's SES. The last two digits of the year the variable is measured are added as a suffix where possible.

Also, if there is nonrandom measurement error for the psychological disorder measures over time, estimates in the previous model may be distorted.

Model I was therefore specified with x_3 , x_7 , and x_8 entered in appropriate nonlinear form in Hennepin and with $r_{e_{se7}}$ and $r_{e_{se8}}$ free parameters in both samples (see Figure 1). Also, path q in Model I was now free, since SES in 1965 was entered in nonlinear form. The most notable contrasts between this respecified model and the original are the significant path a in Hennepin ($a^* = -.119$, $p < .01$), representing the possibility of a nonlinear social

causation effect over the 1965 to 1971 period, and the dramatic improvement in fit in both models. There is evidence of some correlated error in both samples: $r_{e_{se7}} = .431$ in Hennepin and $r_{e_{se8}} = .481$ in Watseka. As a result, the fit of both models is now acceptable. In Hennepin, the probability level of the model is .330; in Watseka, the probability of the model is .885.

If social undesirability effects are the source of the nonrandom error in the items, it is possible this Methods factor is related to SES. Although this covariance is not identified in the present model, it

was possible to specify a Methods factor underlying the substantial error correlations and fix the covariance between this Methods factor and SES in 1960 at some arbitrary value. When this covariance was fixed at $-.3$ in the linear versions of both models, the fit deteriorated in both cases. Thus, the assumption that this covariance is zero is the most reasonable in these data.

Sex and age effects. Sex differences could be confounding the relationship between SES and disorder if women tended to have more symptoms and had lower SES on the average. Also, if age is positively related both to SES and disorder, its absence from the model could be inhibiting possible selection evidence. When sex and age are specified sequentially as x_1 in Model I, it is clear that although each is related to disorder the effects are additive and do not interfere with the interpretation of the social causation or social selection paths in the model.

Across the two samples investigated so far, there is more evidence for social causation than social selection. Attempts to consolidate results for the two regions were unsuccessful, and we are led to conclude that there are real differences between the Hennepin and Watseka samples concerning causal processes for the relationship between socioeconomic status and psychological disorder.

Lenawee County

Analysis of the causal issues in the Lenawee County data is simplified by the fact that there is no variance in sex or age in this sample. All of the respondents are male and 32 or 33 years old in 1972. Table 4 shows the information necessary to derive covariance input for the analysis which follows. A potential problem with the highly attenuated version of the psychological symptom subscale is suggested by the pattern of correlations. Clearly, this scale and the psychophysiological scale are very differently related to other variables in the model. This will result in inordinately bad fit when Model II is estimated as specified in Figure 2. The question is how to decide the relative validity of the two scales as measures of psychological disorder in 1972. It was decided that the more valid measure of disorder should correlate more highly with the 1957 measures which most directly tap the conceptual domain denoted by our theoretical definition of psychological disorder. Two of the 1957 scales qualify as more central indicators of anxiety/depression symptomatology: the nervous symptoms scale from the CTP (x_5), and the nervous tension scale from the 16 P.F. Test (x_7). Table 4 shows that the psychophysiological symptoms scale is more correlated with these measures. Given a forced choice be-

Table 4. Correlations and Standard Deviations for the Cause-or-Consequence Analysis in the Lenawee County Data (N=291)

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
x_1 Father's SEI	...								
x_2 First Job SEI	.293	...							
x_3 SEI (1972)	.254	.733	...						
x_4 WITH (1957)-CTP ^a	-.055	-.142	-.108	...					
x_5 NERV (1957)-CTP	-.048	-.106	-.127	.604	...				
x_6 ANXIN (1957)- 16 P.F.	-.029	-.199	-.112	.381	.297	...			
x_7 NERVENS (1957)- 16 P.F.	.007	-.135	-.086	.409	.509	.385	...		
x_8 PSYCH (1972)	.021	-.009	.024	.223	.211	.167	.212	...	
x_9 PHYS2 (1972) ^b	-.101	-.151	-.150	.195	.228	.119	.244	.496	...
S.D.	22.43	26.38	25.20	3.57	3.24	5.36	4.61	1.04	.929

^a WITH=withdrawal tendencies, CTP; NERV=nervous symptoms, CTP; ANXIN=anxious insecurity, 16 P.F.; NERVENS=nervous tension, 16 P.F.

^b PHYS2=the attenuated psychophysiological symptom subscale from the Langner Index in Lenawee County.

tween them to retain reasonable fit, the psychological symptom scale was removed from the model. Since psychological disorder in 1972 was left then with one indicator, it was necessary to assume that it was perfectly measured (i.e., λ_{p2} was fixed at 1 and $V(e_p)$ was fixed at 0).

Model II was specified with the loadings for all SEI measures assumed equal but unknown.¹⁷ Also, restrictions on the effects of father's SES were necessary due to collinearity problems. The results in the final column of Table 3 show a significant social causation effect from first job status to psychological disorder in 1972 ($a^* = -.120$). Despite the fact that this is early in the individual's occupational career, the effects of low SES are still sufficiently debilitating to increase the chances for the development of anxiety/depression symptomatology up to fifteen years later. The long-term social selection effect is still zero in this model ($b^* = .011$), but there is a short-range selection effect from disorder in 1957 to first job ($w^* = -.185$). This selection effect should not be overinterpreted, however, since there is no chance for occupational mobility over this period. Also, of course, the 1957 measures of psychological disorder are best considered proxy indicators.

A balanced interpretation of this model should include the fact that there is an indirect selection effect of psychological disorder in 1957 on SES in 1972 via first job. The compound path $w^* \cdot c^*$ is $-.168$; thus, whether through the inhibition of educational attainment or occupation mobility, disorder symptomatology does affect the occupational attainments of Lenawee County men fifteen years later.

Nonlinear and correlated error models. Despite the fact that the fit of the linear model is reasonably good (prob. = .0312), it

is still possible that the full power of the effects that have surfaced could better be captured by a nonlinear specification of Model II. A number of nonlinear versions of paths a and b were inspected—mostly by entering various combinations of the SES and disorder indicators in log form. None of these models produced results which deviated from those in the previous section. The possibility that within-test correlated error for the 1957 disorder measures could be affecting the model was also tested by estimating a model in which $r_{e_4e_5}$ and $r_{e_6e_7}$ were free. Both correlations were significant ($r_{e_4e_5} = .25$ and $r_{e_6e_7} = .37$), and the fit of the model improved considerably (prob. > .05), but again the essential results concerning social causation and social selection were unchanged.

The overall conclusion is that some evidence for both social causation and social selection exists in the Lenawee County sample. In view of the ambiguities surrounding the effect of psychological disorder on first job, we might conclude more specifically that evidence for social causation is on firmer ground. When the results for Lenawee County are compared with results for the Hennepin and Watseka samples, it seems that the causal trends in this sample are more similar to the Hennepin sample than the Watseka sample. In both Lenawee County and Hennepin, evidence for social causation exists, while in Watseka no such evidence exists. The overall picture supports the prediction that social causation should play a more central role in explaining negative covariance between SES and psychological disorder. What is perhaps as interesting is the fact that the relevance of social causation seems to change across the three samples.

DISCUSSION

The fact that there is any evidence for social causation in this study is notable in view of the relatively rural samples used. It is entirely possible that the negative covariance between SES and disorder would be stronger in more urbanized samples (Kohn, 1968; Eaton, 1974). The strength of social causation in more heterogeneous samples may in fact be

¹⁷ In Model II it was necessary to set $\lambda_1 = \lambda_2$ in Figure 2 in order to get some estimate of the unreliability in measuring SES in these data. When a variable has only one indicator, three successive measures are necessary before the lambda (λ) parameter can be identified. Although the specification used lumps together background status with the respondent's own status, it was felt that this was more realistic than assuming all the SES variables were perfectly measured.

stronger, and the chances for detecting social selection processes may increase as well. Certainly, there is no reason to expect the social causation evidence found here to disappear either in larger, national panel samples or complementary, urban-based panel samples.

While the panel data from Illinois and Michigan in general support social causation, it should be remembered that these findings are specific to the type of symptoms measured. Judging from the fact that the four-year stability of disorder was .681 in Hennepin and .813 in Watseka (see Table 3), the symptomatology measured in this study cannot be described as either transient or mild.¹⁸ The considerable stability of these symptoms implies that they are something more than temporary responses to recent life stresses or simple indicators of unhappiness. The findings concerning social causation apparently apply to a range of moderately severe, anxiety/depression symptoms which persist over time and thus may be relevant to the diagnosis of a number of nonpsychotic disorders.

In estimating the cause-or-consequence models in this study, it became clear that the results for the Lenawee County and Hennepin models were essentially the same while the Watseka models were quite different. And yet, the Hennepin and Watseka samples are similar in many respects. The task remaining is to engage in some speculation as to reasons for differences in results across the samples. One possible consequence of this is a rudimentary statement of an interaction model for the relevance of social causation in explaining psychological disorder.

Three hypotheses which could explain differences in findings across samples will be considered (although of course others are possible). Differences in the relevance of social causation are explained by:

- (1) differences between samples along a rural-urban continuum;
- (2) the contrasting age/sex composition of the samples;
- (3) differences in the socioeconomic environment of the three samples.

¹⁸ The stability coefficient in Lenawee County is less interesting because the measures of disorder in 1957 and 1972 are different.

Rural-urban differences. This explanation is suggested by Kohn (1968) as a way of explaining the changing nature of the relationship between SES and schizophrenia as the setting moves from rural to urban areas. Although the predicted interaction is specifically relevant to schizophrenia, it is possible the same type of interaction could be operating here. The translation of the prediction to the present case is that as we move from rural (Hennepin and Watseka) to more urbanized (Lenawee County) samples, the relevance of social causation should increase. These samples do not tap the full range of rural-urban environments; the question is whether the range that is tapped is sufficient for the predicted interaction to appear. The presumption is that, for instance, low status in urban areas has more debilitating and fewer manipulable consequences. If this interaction held here, social causation would be operative in Lenawee County and relatively absent in both Hennepin and Watseka. This is not the case: there are clear differences in the importance of social causation in Hennepin vs. Watseka.

Age/sex differences. The prediction in this case is that, controlling for sex, social causation may be more relevant early in the occupational career. Increased stability in extraoccupational roles in later life could buffer the effects of low SES. To investigate this possibility, a cause-or-consequence analysis was conducted on the subsamples of the Hennepin and Watseka data which were most similar in age/sex composition to the Lenawee County data. That is, social causation was investigated on a sample of males aged 26-40 in 1971 in Hennepin ($N = 96$) and Watseka ($N = 63$).

Results showed that in both Hennepin and Watseka a social causation effect appeared over the 1960-1967 lag in Model I. SES in 1960 negatively predicted disorder in both samples, but now with greater strength in Watseka ($g = -.185$ in Hennepin and $-.397$ in Watseka). This gives the social causation by age interaction more plausibility than the previous hypothesis. However, there is still an anomaly. Note that the Hennepin social causation effect in this younger male subsample is not as

strong as in the whole sample. This suggests that social causation is still present in the older part of the sample.

Socioeconomic environment. This explanation is the most logically consistent with our data, but it is emphasized here primarily because the implications for theory are potentially more interesting. The prediction is that social causation becomes more relevant as the socioeconomic environment (or some individual level variable highly related to it) becomes increasingly advantageous and provides opportunities for status advancement. Thus, the more the socioeconomic environment is active, expanding, and

perceived as progressing rather than stagnating, the more that low status probably is experienced as a have-not position in the system.

The differences between samples in socioeconomic environment are represented in Figure 3 in terms of the average rate of increase in SES from first job to current job (in SEI units). Besides the data for the samples in this study, two other groups are added for comparison purposes: (1) a sample of nonblack men ($N = 20,469$), aged 20-64, in the experienced civilian labor force in March 1973 in the Occupational Changes in a Generation-II study for the United States; and (2) a sub-

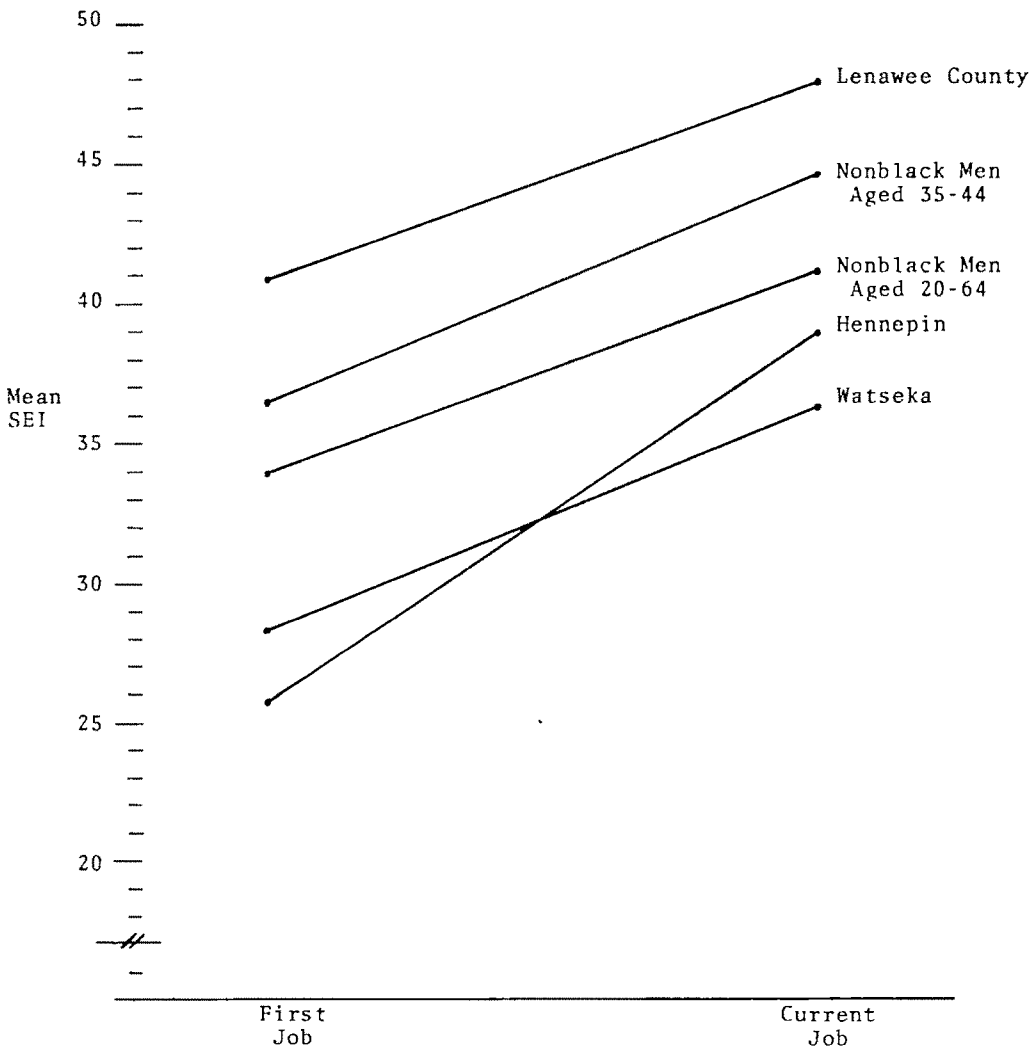


Figure 3. SEI Mobility from First to Current Job in the Three Samples in This Study and Selected Comparison Groups From OCG-II

sample of this same group, age range 35–44 only ($N=4,451$), for comparisons with the Lenawee County sample.¹⁹ These two groups provide the closest baseline information available.²⁰

Figure 3 makes clear that in Hennepin the average rate of status gain from first to current job is higher than in any of the other groups. The average increase of 13 points in Hennepin contrasts with increases of 7.9, 7.1, 7.3, and 8.1 points in the other groups. This is consistent with the fact that the Hennepin area was undergoing industrial development over the 1967–1971 period. Despite the fact that the Hennepin area exhibits a high rate of status gain, it has not caught up to the absolute levels of SES in the national sample. Watseka, on the other hand, is disadvantaged in a double sense: the rate of status gain is only average and the abso-

lute status levels are low. Lenawee County shows only average rates of gain, but sample members start with higher SEI scores and end with higher scores. The data in Figure 3 lead to the following conclusion: the Lenawee County sample is advantaged in terms of absolute levels of status, the Hennepin sample is advantaged in terms of rates of status gain, and the Watseka sample can make neither of these claims. The prediction which is consistent with the data in this study is that social causation will be manifest (Lenawee County and Hennepin) when the socioeconomic environment shows opportunities for either status level or status gain advantages and will not be manifest (Watska) when absolute status levels are low and the atmosphere of development is absent.

This discussion can only serve as a beginning in specifying the situational limits of a social causation effect. But further specifications of sociogenic models of psychological disorder may benefit from the inclusion of variables which interact with SES and which denote the fact that status effects may be stronger at younger adult ages and/or in socioeconomically progressive environments.

APPENDIX

The Four Subgroups of Symptoms from the Langner Index, According to Crandell and Döhrenwend (1967)

Psychological

- B. Periods of days, weeks, months, when I can't get going.
- C. Low spirits most of the time.
- G. Periods of great restlessness.
- H. Worrying type.
- J. Nervousness.
- L. Trouble sleeping.
- N. Memory not all right.
- S. Feel somewhat apart even among friends.
- T. Nothing ever turns out right.
- V. Wonder if anything is worthwhile anymore.

Psychophysiological

- *A. Feel weak all over much of the time.
- *D. Suddenly feel hot all over.
- *O. Bothered by cold sweats.
- *R. Personal worries that get me down physically.
- *V. Frequent headaches.

¹⁹ I would like to thank David L. Featherman for providing the data relevant to the comparisons we are interested in here.

²⁰ Use of an SEI metric could be problematic, since farmers receive an unusually low prestige score on this scale and there are more farmers in both of the Illinois samples than in the national OCG-II sample. Farmers (including owners, managers, and laborers) comprised 4.1% of the current jobs and 11.4% of the first jobs of the OCG-II sample. The comparable figures for current and first job, respectively, in the Illinois data are: 7.5% and 15.9% in the Hennepin sample, and 11.5% and 20.4% in the Watseka sample. To assess the potential for distortion in the comparisons in Figure 3, the mean SEI scores in Hennepin and Watseka were recalculated with farmers weighted in the same proportion as in national data. The resulting means in Hennepin were 26.3 for first job and 39.8 for current job. In the total Hennepin sample, the difference between first and current SEI is 13 points; here, it is 13.5 points. The recalculated means in Watseka were 31.0 for first job and 38.5 for current job. In the total Watseka sample, the SEI difference is 7.9 points; here it is 7.5 points. The difference between Hennepin and Watseka in the rate of gain in SES shown in Figure 3 is preserved when using the weighted subsamples; in fact, the difference is slightly greater. Also, the absolute amounts of gain in the weighted subsamples are very close to the amounts in the total samples. As a result, comparisons between the Illinois samples and the national sample concerning relative change in SEI will not be affected either. The recalculated means in the weighted Illinois samples are all slightly higher than are the means from the unweighted samples (see Figure 3), but the rank order of the means for both first job and current job remains unchanged. Apparently, the comparisons represented in Figure 3 are essentially unaffected by the higher proportions of farmers in the Illinois data.

Physiological

- *F. Poor appetite.
- *K. Frequent fainting.
- *Q. Fullness or clogging in head much of the time.

Ambiguous

- *E. Heart beats hard often.
- *I. Shortness of breath without exercise.
- *M. Acid stomach several times a week.
- *P. Bothered by trembling hands often.

* Judged psychophysiological using the diagnostic manual of the American Psychiatric Association.

REFERENCES

- Bohrnstedt, George W.
1969 "Some observations on the measurement of change." Pp. 113-33 in E. F. Borgatta and G. W. Bohrnstedt (eds.), *Sociological Methodology 1969*. San Francisco: Jossey-Bass.
- Cattell, R. B., D. R. Saunders, and G. Stice
n.d. *Handbook for the Sixteen Personality Factor Questionnaire*. Champaign: Institute for Personality and Ability Testing.
- Clark, Robert E.
1948 "The relationship of schizophrenia to occupational income and occupational prestige." *American Sociological Review* 13:325-30.
- Coleman, James C.
1964 *Abnormal Psychology and Modern Life*. Glenview: Scott, Foresman.
- Crandell, Dewitt L., and Bruce P. Dohrenwend
1967 "Some relations among psychiatric symptoms, organic illness, and social class." *American Journal of Psychiatry* 123:1527-38.
- Dohrenwend, Bruce P.
1975 "Sociocultural and social-psychological factors in the genesis of mental disorders." *Journal of Health and Social Behavior* 16:365-92.
- Dohrenwend, Bruce P., Edwin T. Chin-Shong, G. Egri, F. S. Mendelsohn, and J. Stokes
1970 "Measures of psychiatric disorder in contrasting class and ethnic groups." Pp. 159-202 in E. H. Hare and J. K. Wing (eds.), *Psychiatric Epidemiology*. London: Oxford University Press.
- Dohrenwend, Bruce P., and Dewitt L. Crandell
1970 "Psychiatric symptoms in community, clinic, and mental hospital groups." *American Journal of Psychiatry* 126:1611-21.
- Dohrenwend, Bruce P., and Barbara Snell Dohrenwend
1969 *Social Status and Psychological Disorder*. New York: Wiley.
- Duncan, Otis Dudley
1969 "Some linear models for two-wave, two-variable panel analysis." *Psychological Bulletin* 72:177-82.
- Dunham, H. Warren
1965 *Community and Schizophrenia: An Epidemiological Analysis*. Detroit: Wayne State University Press.
1969 "Social class and mental disorder." *British Journal of Social Psychiatry* 4:76-83.
- Eaton, William W., Jr.
1974 "Residence, social class, and schizophrenia." *Journal of Health and Social Behavior* 15:289-99.
- Fairs, R. E. L., and H. W. Dunham
1939 *Mental Disorders in Urban Areas: An Ecological Study of Schizophrenia and Other Psychoses*. Chicago: University of Chicago Press.
- Fried, Marc
1975 "Social differences in mental health." Pp. 135-92 in J. Kosa and I. K. Zola (eds.), *Poverty and Health: A Sociological Analysis*. Rev. ed. Cambridge, Ma.: Harvard University Press.
- Goldberg, E. M., and S. I. Morrison
1963 "Schizophrenia and social class." *British Journal of Psychiatry* 109:785-802.
- Gurin, G., J. Veroff, and S. Feld
1960 *Americans View Their Mental Health*. New York: Basic Books.
- Haberman, Paul W.
1965 "An analysis of retest scores for an index of psychophysiological disturbance." *Journal of Health and Human Behavior* 6:257-60.
- Haller, Archibald O., and Shailer Thomas
1962 "Personality correlates of the socioeconomic status of adolescent males." *Sociometry* 25:390-404.
- Heise, David R.
1970 "Causal inference from panel data." Pp. 3-27 in E. F. Borgatta and G. W. Bohrnstedt (eds.), *Sociological Methodology 1970*. San Francisco: Jossey-Bass.
- Hollingshead, August B., and Fredrich C. Redlich
1958 *Social Class and Mental Illness*. New York: Wiley.
- Jarrett, William H.
1961 *Personal Consequences of the Stress of Achievement in American Society*. Ph.D. dissertation, Department of Sociology and Anthropology, Michigan State University.
- Jöreskog, Karl G.
1973 "A general method for estimating a linear structural equation system." Pp. 85-112 in A. S. Goldberger and O. D. Duncan (eds.), *Structural Equation Models in the Social Sciences*. New York: Seminar Press.
1976 "Structural equation models in the social sciences: specification, estimation, and testing." Department of Statistics, University of Uppsala. Invited paper for the Symposium on Applications of Statistics, Dayton.
- Jöreskog, Karl G., and Dag Sörbom
1977 "Statistical models and methods for the analysis of longitudinal data." Pp. 285-325 in D. J. Aigner and A. S. Goldberger (eds.), *Latent Variables in Socioeconomic Models*. Amsterdam: North-Holland.

- Jöreskog, Karl G., and Marielle van Thillo
1973 "LISREL—a general computer program for estimating a linear structural equation system involving multiple indicators of unmeasured variables." Research report 73-5. Department of Statistics, Uppsala University.
- Kohn, Melvin L.
1968 "Social class and schizophrenia: a critical review." Pp. 155-73 in D. Rosenthal and S. Kety (eds.), *The Transmission of Schizophrenia*. Oxford: Pergamon.
- 1972 "Class, family, and schizophrenia." *Social Forces* 50:295-302.
- Langner, Thomas S.
1962 "A twenty-two item screening score of psychiatric symptoms indicating impairment." *Journal of Health and Human Behavior* 3:269-76.
- Langner, Thomas S., and Stanley T. Michael
1963 *Life, Stress and Mental Health*. New York: Free Press.
- Lystad, Mary H.
1957 "Social mobility among selected groups of schizophrenic patients." *American Sociological Review* 22:288-92.
- Manis, Jerome G., Milton J. Brawser, Chester L. Hunt, and Leonard C. Kercher
1963 "Validating a mental health scale." *American Sociological Review* 28:108-16.
- Mishler, Eliot G., and Norman A. Scotch
1963 "Sociocultural factors in the epidemiology of schizophrenia." *Psychiatry* 26:315-43.
- Myers, Jerome K., and Lee L. Bean
1968 *A Decade Later: A Follow-Up of Social Class and Mental Illness*. New York: Wiley.
- Myers, Jerome K., Jacob J. Lindenthal, and Max P. Pepper
1974 "Social class, life events, and psychiatric symptoms: a longitudinal study." Pp. 191-205 in Barbara S. Dohrenwend and Bruce P. Dohrenwend (eds.), *Stressful Life Events: Their Nature and Effects*. New York: Wiley.
- Otto, Luther B., and David L. Featherman
1975 "Social structural and psychological antecedents of self-estrangement and powerlessness." *American Sociological Review* 40:701-19.
- Parker, S., and R. J. Kleiner
1966 *Mental Illness in the Urban Negro Community*. New York: Free Press.
- Pelz, Donald C., and Robert A. Lew
1970 "Heise's causal model applied." Pp. 28-37 in E. F. Borgatta and G. W. Bohrnstedt (eds.), *Sociological Methodology 1970*. San Francisco: Jossey-Bass.
- Phillips, Derek L., and Kevin J. Clancy
1970 "Response biases in field studies of mental illness." *American Sociological Review* 35: 503-15.
- Pollin, William, Martin G. Allen, Axel Hoffer, James R. Stabenau, and Zdenek Hrubec
1969 "Psychopathology in 15,909 pairs of veteran twins: evidence for a genetic factor in the pathogenesis of schizophrenia and its relative absence in psychoneurosis." *American Journal of Psychiatry* 126:597-609.
- Robins, Lee N.
1969 "Social correlates of psychiatric disorders: can we tell causes from consequences?" *Journal of Health and Social Behavior* 10: 95-105.
- Rosenthal, David
1970 *Genetic Theory and Abnormal Behavior*. New York: McGraw-Hill.
- Rushing, William A.
1969 "Two patterns in the relationship between social class and mental hospitalization." *American Sociological Review* 34:533-41.
- Schwartz, Carol C., Jerome K. Myers, and Boris M. Astrachan
1973 "Comparing three measures of mental status: a note on the validity of estimates of psychological disorder in the community." *Journal of Health and Social Behavior* 14:265-73.
- Seiler, Lauren H.
1973 "The 22-item scale used in field studies of mental illness: a question of method, a question of substance, and a question of theory." *Journal of Health and Social Behavior* 14:252-64.
- Sewell, William H., and A. O. Haller
1959 "Factors in the relationship between social status and the personality adjustment of the child." *American Sociological Review* 24:511-20.
- Shader, Richard I., Michael H. Ebert, and Jerold S. Harmatz
1971 "Langner's psychiatric impairment scale: a short screening device." *American Journal of Psychiatry* 128:596-600.
- Srole, Leo, Thomas S. Langner, Stanley T. Michael, Marvin K. Opler, and Thomas A. C. Rennie
1962 *Mental Health in the Metropolis: The Midtown Manhattan Study, Vol. 1*. New York: McGraw-Hill.
- Turner, R. Jay, and Morton O. Wagenfeld
1967 "Occupational mobility and schizophrenia: an assessment of the social causation and social selection hypotheses." *American Sociological Review* 32:104-13.
- van Thillo, Marielle, and Karl G. Jöreskog
1970. *SIFASP: A General Computer Program for Simultaneous Factor Analysis in Several Populations*. Princeton: Educational Testing Service.
- Wilkinson, Gregg S.
1975 "Patient-audience social status and the social construction of psychiatric disorders: toward a differential frame of reference hypothesis." *Journal of Health and Social Behavior* 16:28-38.

ORGANIZATIONAL CHANGE IN THE AMERICAN FOREIGN SERVICE, 1925-1965: THE UTILITY OF COHORT ANALYSIS*

THEODORE L. REED

Temple University

American Sociological Review 1978, Vol. 43 (June):404-421

The paper demonstrates the utility of a demographic perspective for the study of organizational reproduction and change; it develops three propositions for the application of cohort analysis to the study of organizational political economy. An analytic perspective on bureaucratic career systems is developed to help account for the response of the American Foreign Service and its elite Officer Corps to the major changes in the international environment during and after World War II. First, organizational lag is traced to the difficulty of changing the demographic composition of the Foreign Service through the maturation of junior officer cohorts. Second, the progress of individual careers (and their organizational consequences) is conceptualized as a function of the relative sizes and locations of entering cohorts. Finally, the importance of control over cohort succession is seen through an examination of the complex patterns of cohort transformation (promotions) during the postwar years. Analysis of the data reveals a variety of practices which buffer the Foreign Service from pressures for rapid transformation. The summary identifies conditions under which studies in the political economy of organizational reproduction and change will benefit from the examination of cohort processes.

Recent literature on organizations has made it increasingly clear that continuity and change are not analytically distinct social phenomena (Zald, 1970a; 1970b; Benson, 1975; Hinings et al., 1974; Perrow, 1972; Thompson, 1967). The persistence of institutions despite exogenous and internal pressure for their transformation provides an important opportunity to study the determinants of organizational reproduction (Hernes, 1976:514).

This paper develops a middle-range, theoretical explanation of bureaucratic persistence and change that identifies the constraining influence of cohort organization. More specifically, I use a cohort perspective to link empirical work on adaptation by the American Foreign Service bureaucracy to changes in the international environment during and immediately after World War II.

Three general propositions utilizing this perspective are developed through an examination of the relationship between bureaucratic change and demographic structure in the post World War II Foreign Service Officer Corps. Data for the first

two are drawn from existing literature. The operations described in the third are used to help interpret data gathered on Foreign Service promotions. The propositions, which are used as guideposts for the discussion to follow, are:

(1) To the extent that organizational adaptation requires changes in composition of personnel skills and attitudes, cohort organization will contribute to a lag in organizational responsiveness. Earlier research has recognized that throughout a period of significant change in organizational task environment, the Foreign Service remained remarkably stable, exhibiting little of the expected structural adaptation. By maintaining a career in which senior posts were reserved for older and well-socialized officers, the Foreign Service Administration eschewed newer missions and innovative techniques demanded by modern diplomacy (Brookings Institution, 1951; Committee on Foreign Affairs Personnel, 1961; Harr, 1969).

(2) The distribution of organizational rewards, as reflected rate of promotion for organizational members, is influenced by the relative sizes and placement of cohorts. There is already a body of evidence which suggests that slowed promotion rates, particularly among junior For-

* I am indebted to Leo Rigsby, Marilyn Lester, John Houghton, David Elesh and two anonymous referees for helpful criticism of earlier drafts.

eign Service Officers, has been an important roadblock for those who assumed that rapid promotion would enable them to apply new techniques to the unique problems of modern diplomacy (Martin, 1968; Harr, 1969; Commission on the Organization of the Government for the Conduct of Foreign Policy, 1975).

(3) The maintenance of a particular cohort career form and a preferred demographic distribution has important consequences for the organizational political economy. The preservation of the traditional structure of the Foreign Service Officer Corps meant that younger officers whose social origins were more representative of the public at large and who shared a newer conception of the role of the Foreign Service consistently were underrepresented among senior decision makers in the Department of State up through the decade of the 1960s (Segal and Willick, 1968; Harr, 1969).

THEORETICAL SETTING AND ORGANIZATIONAL CONTEXT

A secondary review of existing studies and an extensive analysis of longitudinal data on changes in the Foreign Service reveal the linkages between these discrete observations and demonstrate the complex but important relationship between the demographic composition of the organization and the broader political and economic features affecting bureaucratic persistence and change. The concept of cohort provides a major interpretive tool for the analysis—the subject reviewed next. That is followed by a discussion of the important features of the Foreign Service, which sets the stage for theoretical development.

A. Cohort Analysis

A cohort refers to a population aggregate experiencing the same events during the same period in its lifetime or, in this case, in its career (Ryder, 1965:845). Cohort analyses can have at least three major foci. First, as society changes, each new cohort encounters a unique sequence of social and environmental events. That is, macrosocietal changes affect cohort experiences; hence, the life-course pattern of people in one cohort will differ in some respects from the life-course patterns of those in others (Riley, 1973:36). In this vein, cohort analysis directs attention to the effects of cohorts emergent at one point in time, on later organizational experiences. In organizations cohort differences are associated with changes in training (e.g., a new graduate curriculum), experiences of personnel (e.g., the unusually favorable job market and grant possibilities facing new Ph.D.s in the middle and late 1960s), and with changes in organizational goals (e.g., the different objectives of students in the 1960s vs. 1970s). The cohort approach directs attention, not to a summation of individual life or career histories, but to the distinctive composition of successive cohorts.

The comparative study of cohorts is, then, an inquiry into "structural transformation rather than the network of actions and interactions predicated in the routine operation of institutional structure" (Ryder, 1965:843). This analysis of structural transformation will help to explain the interrelated character of organizational careers among successive cohorts of Foreign Service Officers, hereafter FSOs, and changes in the structure of the Foreign Service Administration

Table 1. Average Promotion Indices (PI) for Five-Year Intervals for Cohorts Completing Twenty Years of Service

	Cohort				
	1925	1930	1935	1940	1945
5 years	55.6	00.0	24.2	102.3	357.4
10 years	135.6	126.0	282.2	345.7	77.3
15 years	162.5	209.3	287.7	83.3	100.0
20 years	150.6	240.4	95.8	136.0	54.6
Average	126.1	143.9	172.5	166.8	147.1
	(18)	(32)	(16)	(23)	(19)

and its immediate power setting (Warwick et al., 1975). Both continuity and changes in cohort composition reflect the application of recruitment and promotion rules developed by Foreign Service administrators.

Second, the cohort (or generational) organization of society has a direct bearing on the conditions for change, i.e., is an independent variable in social change. Major structural transformations of society have frequently been associated with the unique experiences, conditions, problems and opportunities facing particular generations (Mannheim, 1952; Stinchcombe, 1965; Ryder, 1965; Zeitlin, 1970; Inglehart, 1971; Nagle, 1975). Shared cohort experiences, rather than state-party bureaucratization or social class formation are seen as the basis for the character of the post-1939 Soviet elite in Nagle's analysis. Zeitlin uses a similar generational model to explain differences in the level of support of the Cuban revolution.

Third, organizational change is closely linked with variations in the demographic composition of particular and successive cohorts (Hannan and Freeman, 1977). Everyday examples are legion. The growth of American universities in the 1960s is in part attributable to maturation of the post-World War II baby boom cohort. The expansion of graduate cohorts during those years has been cited as a reason for the oversupply of academic talent during the current period of university retrenchment. Promotion opportunities likewise have reflected the changing relationship between cohort size and the number of available positions. Anthony Downs (1967) has concluded that the responsiveness of federal bureaucracies is heavily dependent on the age and seniority distribution of personnel. This finding confirmed Lipset's observation that political change was impeded by senior cohorts of bureaucrats in Saskatchewan (Lipset, 1959; 1960). Michels's (1958) iron law of oligarchy can be identified as another specific case of this general phenomenon, in that oligarchies are frequently cohort based.¹

¹ More recent research, which has developed the

B. *The Foreign Service*

The Foreign Service is a heteronomous organization of professionals within the larger Department of State complex, and, like the Civil Service, it is a semi-autonomous personnel system within a larger bureaucratic structure (Downs, 1967:247; Heydebrand, 1975:13). The term Foreign Service traditionally has two connotations: (1) the Foreign Service Officer Corps, and (2) the administrative body.

(1) Individual members of the Foreign Service are called Foreign Service Officers or FSOs; collectively they are identified as the Officer Corps. As a body, the Officer Corps is mandated to oversee much of the administration, formulation, and actual execution of American foreign policy. While assigned primarily to Foreign Service posts overseas, a large group of officers regularly serves in the Department of State in Washington. Historically, most FSOs have entered the Service at junior officer levels, after a rigorous examination process.² Unlike Civil Service employees, whose ranks are attached to position, FSOs are assigned a personal rank similar to the pattern found in the military. Through a process of annual evaluation and review, officers are promoted through the ranks until they reach the senior grades (FSO-1 or FSO-2) or until their careers are interrupted by separation or retirement. Beyond that, a supergrade of Career Minister or Career Ambassador identifies an FSO as one of a select group included in the ranks of the senior decision makers within the American foreign policy establishment.

The prototypical FSO career goes through four stages associated with increased responsibility and promotion through the ranks of the Officer Corps. Incoming junior officers go through a probationary and training stage for several

cohort perspective more fully, is reviewed in the discussion of the three propositions.

² With the exception of a group of lateral entrants in the early 1950s and scattered individuals who have joined the Service at the midcareer level, officers begin their careers as junior officers (FSO-8 or FSO-7) and continue until retirement with senior officer status (FSO-1, Career Minister or Career Ambassador).

years with personal ranks of FSO-8 and FSO-7. In midcareer they develop functional and area specialties and proceed from rank FSO-6 or FSO-7 to FSO-4 or FSO-3. In the beginning of the senior stage (ranks FSO-4 to FSO-2), they typically are given broadening assignments and program direction responsibilities. As senior officers at the leadership stage, FSOs are given key positions in the department or overseas involving policy formation or program direction. Senior officers usually hold personal ranks of FSO-2 or FSO-1. The most successful officers hold the personal rank of Career Minister or Career Ambassador, comparable to the General/Admiral ranks in the military. The particular class structure and distribution of all members of the Officer Corps for selected years is reported in the Appendix.³

(2) The term Foreign Service also refers to the administrative body (hereafter referred to as the Foreign Service Administration) which oversees the selection and promotion of FSOs and other

categories of Foreign Service personnel (e.g., Foreign Service Reserve and Foreign Service Staff) not considered in this paper.⁴ The Foreign Service Administration is a subsidiary people processing organization charged under a broad mandate from the larger Department of State bureaucracy to recruit, train, and oversee the careers of these key organizational personnel.

The modern Foreign Service, an amalgam of earlier diplomatic and consular services, was established formally with the passage of the Rogers Act of 1924. At the time of its creation, the members of the Foreign Service, as depicted above, comprised the bulk of civilian overseas representation.

While the Service initially was mandated to provide personnel for all top level civilian positions overseas, there always has been considerable debate about the extent to which such an assignment is either realistic or desired by the Foreign Service itself (Barnes et al., 1961). During the years since its establishment, the volume and complexity of American overseas involvement has increased considerably. This was fostered by the expansion of American political interests, accelerated by the events of World War II, and reflected in an enlarged task domain. Several prestigious investigatory commissions have recommended that the Service expand and recruit specialized personnel especially to encompass new economic (i.e., AID) and information (i.e., USIS) functions. Although it is not as technology-based as the military profession, changes in communication technology has made Washington-based decision making and control more feasible in foreign affairs work (Harr, 1969:324). Together, with the expanded knowledge base, these changes have made the independent diplomat obsolete. There is an increased need for specialists overseas to gather new information and experts at

³ Entering cohorts, attrition and midcareer senior entry all change the size and composition of the Officer Corps. Entering cohort sizes for the sample cohorts are reported in fn. 5. The growth during the years 1947-1965 averaged 123 new officers per year. (The range was from a low of 50 in 1950 to a high of 292 in 1956, with a standard deviation of 69.) Since this growth occurs in cohorts following those studied here, it has little, if any, influence on the patterns of the mobility described herein. Attrition rates computed for the sample cohorts exhibited little inter- or intracohort variability over time. Lateral entry, or the introduction of new officers at midcareer or senior levels was negligible prior to 1956.

The Department, following the recommendations of the Wriston Committee (1954), opened the Foreign Service Officer Corps to transfers from other personnel categories within the Department of State. Between 1954 and 1958, 1,525 officers were added to the Officer Corps in this manner. Despite dire predictions about the introduction of the Wristonnees making promotion to the senior grades more difficult for those who entered at the bottom of the Service prior to 1954, by 1956 less than one-third of 1% of Wristonnees eligible for promotion were promoted compared with a comparable promotion rate of 46% for nonlateral entrants (Steiner, 1958:17,31). Promotion rates were roughly equalized by 1961. While it is clear that this program of lateral entry had widespread effects on the Officer Corps, lateral entrants for the most part entered the Service at ranks below those from the cohorts being studied. For this reason, they did not have any direct influence on promotions of the cohorts studied herein.

⁴ The assignment of officers is not formally a task delegated to the Foreign Service Administration. Although other administrative structures in the Department of State are responsible for decisions on location and for functional assignments, management experts in the Foreign Service have considerable influence in these decisions.

home to supplement their analyses. The Administration of the Foreign Service, during the years under study, reflected the views of older cohorts by concentrating on traditional Foreign Service functions—representation, political reporting, and miscellaneous consular tasks.

The personnel decisions in the Foreign Service are made within a frame of reference which is similar to that found in other bureaucratic subdivisions concerned with the regulation of careers for professional employees (Weber, 1947; Kaufman, 1960; Janowitz, 1960; Hall and Schneider, 1973). Administrative units in many large scale organizations conceptualize their organizational functions in cohort terms. They recruit entering cohorts (usually on the basis of a calendar, fiscal, or academic year) and administer entry level orientation or training to a whole cohort or selected clusters within that cohort. Frequently, they move individuals and manage their careers in response to manpower needs of the organization and the advancement and task expectations of successive cohorts (Kanter, 1976). The Foreign Service is no exception to this general pattern. The similarities are reinforced by defined organizational units and procedures for handling each stage in the maturation of the cohort. These administrative procedures guarantee that seniority is the single most important determinant of promotability from one personal rank (or grade) to the next (Reed, 1972:191–2). Other research suggests that FSOs entering the Service at the same time regard themselves as a cohort and make informal comparisons within the cohort to judge career success (Feinstein, 1969; Scott, 1969).

Despite the usefulness of the cohort perspective to describe groups of FSOs being processed by the Foreign Service Administration, and the recognition of the cohort as a point of reference among FSOs themselves, there is no systematic application of cohort analysis, as developed in demography, to understand the operation of the Officer Corps. A number of propositions relating cohort characteristics and organization processes are developed and explored in the remainder of this paper.

THE LAGGED RESPONSE TO CHANGE

As previously indicated, a number of studies have highlighted the incongruity between the dramatic change in the conduct and complexity of international relations beginning in World War II and the persistence of Foreign Service organization and practices. Changes in the task environment included the addition of new functions requiring new personnel and expertise in information processing (both dissemination and collection); economic analysis and foreign assistance (Committee on Foreign Affairs Personnel, 1962; Macomber, 1970); expansion in the number of overseas missions requiring Foreign Service staffing (Hoover Commission, 1955; Wriston Committee, 1954); and the need for new personnel and techniques to handle the administrative and managerial functions necessary for coordination and control of the expanded operations (Hoover Commission, 1949; Brookings Institution, 1951; Steiner, 1958; Harr, 1969).⁵

Many of the same reports provide extensive documentation of Foreign Service persistence despite these pressures for change (Jones, 1965; Commission on the Organization of Government for the Conduct of Foreign Policy, 1975). Earlier research reports that apart from other determinants of organizational reproduction, the demographic composition of organizations seriously constrains their ability to adapt rapidly to external change (Stinchcombe et al., 1968:22; Tepperman, 1975; Lipset, 1959). These works provide evidence to support the operation of the first proposition, as presented at the beginning of this paper. To understand how an existing method of cohort organization constrains adaptation however, it is necessary to examine the issues in greater detail.

The interval between the reception of

⁵ Changing organizational needs were also a consequence of growth in the number of Foreign Service posts staffed by members of the Officer Corps. The United States maintained diplomatic relations with 62 countries in 1939, 79 countries by 1950, and 87 countries by 1960. The number of major posts (Embassies, Legations, Consulates General, Missions of Special Offices) grew as follows: 1936, 91; 1940, 99; 1945, 101; 1950, 126; 1955, 140; 1960, 167; 1965, 189.

information concerning the changed conditions in the environment and the execution of the corresponding step in the goal-seeking behavior of the system has been identified in the system's model of organizations as a time lag (Deutsch, 1963:187-8). It is seen primarily as a consequence of the information feedback delays necessitated by reception and processing and has been used to describe routine adjustments to environmental fluctuations. When organizations undergo major structural changes, additional information processing problems contribute to the lag. The collection and use of new types of information frequently require new organizational skills and different perceptions of organizational goals. These adjustments involve major changes in the demographic composition of an organization.

One such time lag in organizational response is revealed by the relationship between the expanded domain of the Department of State during the post-World War II period and changes in manpower composition and functions. This is a situation in which the centrality of the professional career in the existing manpower structure of the Officer Corps delayed the adaptive process and produced an extended time lag.

The Officer Corps, at the end of World War II, operated with a personal ranking system, little lateral entry, and a relatively small (approximately 1,400) body of personnel. The consequence of the closed career service structure, identified by the cohort organization, is that commitments of people at senior and middle levels of the organization are determined by the composition of cohorts recruited much earlier (and under the constraints of different needs). Personnel resources change primarily through modifications in lower level recruitment and training. While the aforementioned pressures for change in the State Department were reflected at all levels of the organization, actual change most readily occurs at the lowest ranks where people with new commitments and skills are recruited. Change at the middle and senior ranks of the Officer Corps depends on the slow filtering of these new recruits through the ranks to higher levels.

Because this closed career system was highly valued, the Foreign Service Administration directed attention to procedures governing size and composition of new entering cohorts, rather than toward bringing about immediate change by recruiting for middle and senior level policy positions from outside the organization.

Examination of the exceptions to this general pattern reveals how the Foreign Service Administration circumvented some of these demographic restrictions. During the emergency situation in World War II, a new temporary personnel category, the Foreign Service Auxiliary, provided more manpower and particular administrative skills necessary to handle complicated consular, commercial, and war-related matters. The auxiliary personnel filled middle and in some cases, senior level positions left vacant (or unfillable) by the traditional procedures. Subsequently, this alternative personnel category, which enabled the organization to respond to new or unique external demands without the ten- to thirty-year lag implied by the maturation of junior cohorts, was re-created on a permanent basis as the Foreign Service Reserve. Reserve Officers were hired on a temporary basis (maximum three-year term) and since the early 1970s on a permanent basis to meet personnel needs outside the scope of Officer Corps competence, but the structure of the traditional Officer Corps remained intact, buffered in part by this organizational device.⁶

The lagged response of the formal Offi-

⁶ The creation of the Auxiliary was not the only alternative developed during this period for coping with the slowness of change in the Officer Corps. The lateral entry of several hundred officers during the Wristonization program of 1954-1958 was another controversial adaptation. Special training was also accelerated. The lag between organizational demands for new or different skills and the ability of the Foreign Service Administration to respond through the personnel structure of the Officer Corps is illustrated by the reaction to introduction of new foreign affairs functions. Although many in the leadership of the Department of State favored expansion into newer techniques in the field of international relations—new forms of intelligence gathering, economic aid, and new information functions—the limits of the Officer Corps adjustment facilitated transfer of many of these functions to new agencies (USIS, AID) or to the military/NSC apparatus.

cer Corps itself, in this instance, is typical of the change process identified in traditional demographic cohort analyses (Ryder, 1965). In addition, the use of new criteria for the selection of entering cohorts has produced challenges to the existing professional self-definitions from junior officers recruited as specialists into an Officer Corps dominated at senior levels by generalists (Harr, 1965b:267-86). Similar phenomena have been identified in university faculties and in the military where personnel composition is a function of long-past recruitment decisions, cohort transformation, and changing organizational context (McNeil and Thompson, 1971:629; Moskos, 1970:178-9; Janowitz, 1960; Kaufman, 1960).

INDIVIDUAL PROMOTIONS AND THEIR COHORT DETERMINANTS

Individual careers, because of their intimate connection with the reward structure of complex organizations, are the subject of intense interest in any large scale bureaucracy or professional network. The Foreign Service is no exception. In any career there are marked individual differences in rates of mobility (dependent on the social origins, training, experience, and relative seniority) and many of these have been identified for the Foreign Service (Reed, 1972; Harr, 1969). But more basically, the variation in career mobility rates is a function of how the system is organized. Generally, the cohort demography of any complex organization influences both the overall shape of the career and its internal stratification (the distribution of individuals). Thus, the second general proposition which guides the discussion below identifies the relative sizes and placement of cohorts as influences on individual career patterns and thereby, on processes of organizational reproduction and change.

Particular cases of this general relationship have been reported in a number of earlier studies. The specific consequences of cohort maturation and the changing demographic composition of an organization were observed in research on the post-World War II military (Janowitz,

1960). Younger officers who rapidly rose to senior positions during the war facilitated organizational renewal. This was traced to the expansion of promotion opportunities for officers during the war which, in turn, was identified with the introduction of large cohorts of enlisted personnel and draftees.

Janowitz, writing about the consequences of cohort maturation processes in the military, anticipated increased similarity between the skills and orientations of military and civilian leaders in the decade of the 1960s as a result of these developments (Janowitz, 1960:15) and identified two sources of this change. First, he described the increasing organizational and technological complexity of warfare. This, he postulated, would make the personnel needs of the military similar to those of other large scale organizations. Furthermore, he recognized that the rapid succession of cohorts would bring younger officers (with new skills) into the senior ranks of the services. In this case, the time lag in change, associated with the promotion of younger cohorts to senior positions, was both recognized and hypothesized to provide the vehicle for significant organizational redirection in the future. But despite some terminological difficulties with Janowitz's conception of convergence, subsequent analyses have found a lack of convergence and have sought alternative explanations (Kanter, 1976).

As the military ceased to grow during peacetime, promotion rates slowed. Large cohorts of relatively senior officers were caught in the bottleneck of restricted promotions at senior ranks, reflecting the degree to which individual rates of promotion depend on differences in the size of entering cohorts and their demographic transformation (i.e., aging within the organization). During the 1960s, although major contractions in the size of the military had occurred more than a decade earlier, mobility was reduced at lower levels because relatively large cohorts remained at middle and upper ranks. Not only was the mobility of these officers slowed, but the situation also prevented the promotion of younger officers with new skills and perspectives. Thus, cohort composition

seriously constrained both individual careers and organizational change.

An important contrast in consequences for personnel arises when we compare organizations with graded positions vs. graded personnel. Mobility under the first condition is a function of vacancies into which one can be promoted, paralleling graded positions. So when the initial position in the vacancy chain is at a relatively senior level, a greater number of position holders is affected than when that initial vacancy is at a junior level (White, 1970). The creation of new senior level positions (vacancies) will accelerate promotion rates for those below. In the Foreign Service, by contrast, grades are personal and not identified precisely with particular vacancies. Unlike many large scale organizations in which there is an immediate and direct connection between individual mobility and fluctuations in the number and location of vacancies, openings almost always exceed available staff in the Officer Corps which reflects the second kind of ranking principle.⁷ For this reason, the volume and distribution of supply is of immediate practical and continued theoretical concern.

Because of dramatic variations in the size of entering cohorts between the mid-1950s and mid-1960s, the Foreign Service Administration faced a new problem of harmonizing individual career expectations (rapid promotion for the able as in the immediate post-World War II period) with the overall interests of the Officer Corps (maintaining a balanced rank distri-

bution of personnel and the assignment of officers having high personal rank to prestigious senior positions).

The individual and collective protests of junior Foreign Service Officers during the late 1960s grew out of a perceived change in the ground rules for promotion which adversely affected those at lower levels. Compared both with older FSOs and with their own counterparts in other organizations, many younger officers felt they were being promoted too slowly (JFSOC, 1967; President's Task Force on Government Organization, 1967; Martin, 1968). The ramifications of this dilemma for the organization were widespread, including the eventual unionization of Foreign Service Officers and the loss of many capable officers through resignation. The lack of the responsiveness (not utilizing newer skills and techniques, not promoting younger officers to positions of responsibility) frustrated many senior policy makers during the Kennedy years and since (Commission on the Organization of Government for the Conduct of Foreign Policy, 1975).

While staffing needs remain vaguely defined, except at top policy levels, longitudinal projections (simulations) are now used by the Foreign Service Administration to estimate staff supply (Commission on the Organization of Government for the Conduct of Foreign Policy, 1975). Like many complex organizations, the Foreign Service has been able to develop a proactive stance because of these efforts (Hickson, et al., 1971).

We have thus seen that some of the personnel problems faced by the Foreign Service Administration and even some of the annoyance with the inflexibility of the Department of State are traceable to the influence of differences in cohort size, maturation, and location on individual FSO careers. The range of promotion opportunities is a function of the size of preceding cohorts and their relationship to vacancies further up the organizational hierarchy. Either decreases in vacancies (both relative and absolute) or relatively large preceding cohorts may slow the progress of individual careers. Changes in the ground rules for careers are an important determinant of the flexibility of organ-

⁷ Traditionally, the Department either has been nominally understaffed or has filled these vacancies with temporary personnel from the Foreign Service Reserve. This practice has two roots. (1) Largely voluntary restrictions on the size of entering cohorts help to explain this manpower shortage. The practice is, of course, familiar to students of the professions. Market scarcity and the employment of nonprofessional assistants help maintain professional control over both the definition of professional work and, consequently, professional rewards. Traditionally, the least desirable Foreign Service positions remain unfilled by full-fledged members of the profession. (2) Because rank in the Foreign Service is associated with person rather than position, there is considerable latitude in the matching of rank and responsibility of assignment. Unlike the military, the Foreign Service does not use brevet ranks.

izational adaptation to environmental change.

THE POLITICAL ECONOMY OF COHORT SUCCESSION

A. Cohort Organization and Organizational Reproduction

Empirical studies of cohort succession within complex bureaucracies have occasionally gone beyond identifying instances of bureaucratic persistence to hint at the existence of forces favoring organizational reproduction. In short, they have recognized that the dynamics of cohort organization necessarily influence succession patterns. For example, Mayer Zald (1970a) has identified the succession system as an important element in the operating political economy of the organization. Arnold Kanter's (1976) recent examination of managerial careers among Air Force generals and Segal and Willick's (1968) analysis of traditional career patterns in agencies under stress compare the skill composition of leadership cohorts over time. Both studies find that while changing cohort composition at lower ranks (e.g., skills, backgrounds, etc.) is important, it is not necessarily a guide to leadership composition because differential intracohort promotion rates tend to favor those with the more traditional orientations. Ralf Dahrendorf (1968) makes a similar point about stratification systems in his essay, "On the Origins of Inequality Among Men." Pressure against rapid regeneration of organizations (McNeil and Thompson, 1971) and the existence of specified goals for preferred demographic distribution (Stinchcombe et al., 1968) both have been cited as causes of similar patterns in other bureaucracies.

These findings are consistent with the basic insights of the political economy perspective on organizations. Any understanding of the consequences of cohort demography for structural change is linked to a pervasive contest for power, privilege and prestige within the organization (Zald, 1970a; Hass and Drabek, 1973; Thompson, 1967). Cohort organization and composition have important consequences for the reproduction of the organ-

izational structure and the larger pattern of social relations into which it is bound (Wright, 1977). From the foregoing literature, I formulated the third proposition: (3) The commitment to a particular cohort form and a preferred demographic distribution become an important tool in the contest for power, privilege, and prestige described by the organizational political economy.

The complexity of processes of organizational reproduction and the relevance of the cohort approach are revealed by the study of changing promotion rates for selected cohorts and years between 1925 and 1965. The general relationship stated in the third proposition guides the search for sources of significant inter- and intra-cohort variation over time. The account below briefly reviews the nature of external demands but focuses primarily on the organizational response of the Foreign Service Administration.

The general direction of external pressures for change is most clearly indicated by the plethora of high level investigatory commissions which recommended increases in the size and the development of new functional specialties for the Officer Corps. Harr (1965b:9-24) reviews their recommendations and their projections of the personnel requirements for a new diplomacy. All agreed that the Foreign Service would require substantial changes in existing manpower resources.

Given the breadth of the proposed changes, any response of the Foreign Service Administration was likely to have important consequences for the demography of the Service and vice versa. Two of these already have been discussed. First, new functions (foreign aid, economic forecasting, military liaisons, etc.) were delayed in their implementation because of the required recruitment and succession of new personnel who incorporated the new skills and views of the Officer Corps mission. Second, changes in the apportionment of personnel throughout the ranks of the career service influenced the mobility of career officers, thereby drawing them into determination of the shape of proposed changes. The political economy of the organization is reflected in other demographic constraints as well. In

this setting, the options for demographic change are influenced further by commitment to existing career structures (e.g., the rank-in-person career system which determined succession patterns) and the priorities of both internal and external decision makers (e.g., the commitment of senior FSOs to maintenance of the traditional diplomatic functions of political reporting and representation as the central mission of the Officer Corps). Both statutory mandate and the Foreign Service Administration prescribed a professional service renewed primarily through recruitment at junior ranks. Within these constraints, numerical expansion was achieved through increases in the size of entering cohorts (see Appendix, fn. c). Extensive documentation indicates that changing functional specialties and demands for a more representative cross section of the population were reflected also in recruitment priorities (Harr, 1969:167-70; Segal and Willick, 1968).

The important influence of organizational political economy for structuring the cohort demography is revealed by adaptation at midcareer and senior levels for these necessitate a more complex series of changes. The critical processes are those governing cohort transformation and succession, in this case measured by promotion rates. The most obvious option, an acceleration in promotion rates, would place younger officers in ambassadorial and ministerial positions. The maintenance of existing rules governing cohort transformation (promotion), on the other hand, would make it extremely difficult to implement the recommended changes.

Pressured into some accommodation, but unwilling to give up the existing elite Officer Corps, the Foreign Service developed other alternatives. Promotion rates (which govern cohort transformation) and the number of years of service required before promotion to senior levels provide reliable and surprisingly direct indicators of the Foreign Service responses. The closed career system in the Officer Corps, of course, facilitates application of insights derived from the cohort approach.⁸

⁸ Although there are officers who enter the Officer

B. The Data on Promotions

In order to study the Foreign Service response to the increased need for FSOs, a sample was drawn from those officer cohorts still in the Service during the period surrounding the rapid changes of the World War II era. All members of a sample of cohorts, those entering the Service in 1925, 1930, 1935, 1940, and 1945, were included in the analysis.⁹ The promotion experience for each of these cohorts was summarized for five-year intervals during the period under investigation. An index of promotion was developed according to the following formula for comparing the career mobility of all cohorts,

$$PI = \left[\sum_{i=1 \text{ to } n_t} (c_{t+5} - c_t) / n_t \right] \times 100,$$

in which PI is the promotion index, c_t is the class/grade of an officer at the beginning of the five-year period, c_{t+5} is the class/grade of an officer at the end of any given five-year interval, and n_t is the size of the cohort at time t .

In the resulting index, a promotion of one rank for every officer in a cohort during a five-year period would yield a rate of 100. Since officers may be promoted more than once during a five-year time interval, it is possible for this index to exceed 100. Although there are no cases in which an officer was demoted, it is, of course, possible for a person not to be promoted within a given interval.¹⁰ Because there is attrition during each five-year period, cohort size on which the promotion index

Corps without going through the examination process, and others who are classified as Reserve or Auxiliary Officers (hired for short periods to perform tasks requiring specialized skills), these two groups are small relative to the size of the Officer Corps. In 1961, long after most of these officers had an opportunity to gain seniority, Harr (1965a) reports that fewer than 20% of the officers of decision-making rank were from other than regular FSO classification.

⁹ Entering cohort sizes for the cohorts studied were as follows: 1925, 36; 1930, 54; 1935, 24; 1940, 41; 1945, 42. Data on promotions and careers were gathered from the Biographic Register of the Department of State.

¹⁰ Although the number of grades in the Service was changed from eight to ten classes in August 1956 and some officers were reduced in grade at this time, the change did not affect those under study in this paper.

was based decreased during successive intervals.¹¹

C. Organizational Reproduction: The Significance of Cohorts

To increase its size and change its composition, particularly at senior levels, the Foreign Service Administration had to increase promotion rates. Promotion experience indices for each of the five cohorts were plotted against the dimension of absolute time (see Figure 1). There is a marked increase in the promotion index during the 1940–1950 and especially during the 1945–1950 time periods, followed by a return to lower promotion rates in the decade of the 1950s. This holds for each group except the 1925 cohort whose surviving members were predominantly at senior levels by the 1940–1950 periods and thus could no longer be promoted. While the Foreign Service appears responsive to pressures for growth during and immediately after the Second World War, the intracohort changes exhibited since 1950 are not as easily interpreted with reference solely to external demands for expansion. Pressures from the external power setting of the Foreign Service during the postwar years were not decisively united behind either expansion or contraction. Accounts of the period reveal that, on the one hand, rapid military growth in response to the intensified cold war and the loyalty investigations of Senator Joseph McCarthy aided Foreign Service retrenchment. Further expansion was halted by decreased reliance on traditional diplomatic vehicles in foreign affairs during the early half of the 1950s.

On the other hand, powerful sectors of the business and political establishment sought Department of State coordination of increased international commerce, expanded foreign aid programs and enlarged information facilities. These changes required both new foreign policy initiatives and an expanded Foreign Service staffed by officers with new orientations and new

skills (Hoover Commission, 1949; Harr, 1969:22).

While this complex matrix of forces shaped the external political economy of the Service, it is difficult to understand how they could directly produce the remarkable *intercohort* differences in promotion which characterize these years. A detailed examination of sampled cohorts reveals that the depths of the postwar decrease in promotion rates is proportional to the magnitude of the increase during the earlier period, a pattern associated with attempts at system reequilibration, not change. Figure 2, which presents the same promotion data using years of service of the cohort rather than an absolute time scale, more clearly reveals the differentials between the experiences of the 1925 and successive cohorts.

Senior officers are more protected from environmentally induced fluctuations. The patterns for older cohorts, and hence of those officers further along in rank, are less affected by the dramatic wartime and postwar events than are those of more recent entrants. As indicated above, the 1925 cohort exhibits a rather smooth curvilinear distribution. Each successive cohort—1930, 1935, 1940, and 1945 displays an increasingly peaked pattern coinciding with the events of the 1945–1950 years (indicated by the final year of this period, 1950, in Figure 2). For each later cohort, members are of lower rank. Thus, promotions to junior and middle ranks during these years are proportionately more frequent than promotions to senior ranks. In addition, the rebound to lower significantly rates of promotion in the 1955–1960 period is correspondingly greater for the same sequence of cohorts. This pattern of adjustment is the behavior of a system which has overreacted in response to environmental threat (1945–1950) and then has engaged in corrective action to reproduce the status quo ante in the promotion rate for officers. Intercohort differences demonstrate that promotions at the higher ranks are less susceptible to environmental pressures than are those at lower ranks.¹² To explain why

¹¹ Cumulative attrition rates for each cohort, over time, are presented in Figure 3. A distribution of ranks for FSOs completing 20 years of service reveals that promotion ceilings are also unlikely to produce the observed intercohort differences.

¹² Neither the work of the systems theorists nor the more recent formulation of Hass and Drabek,

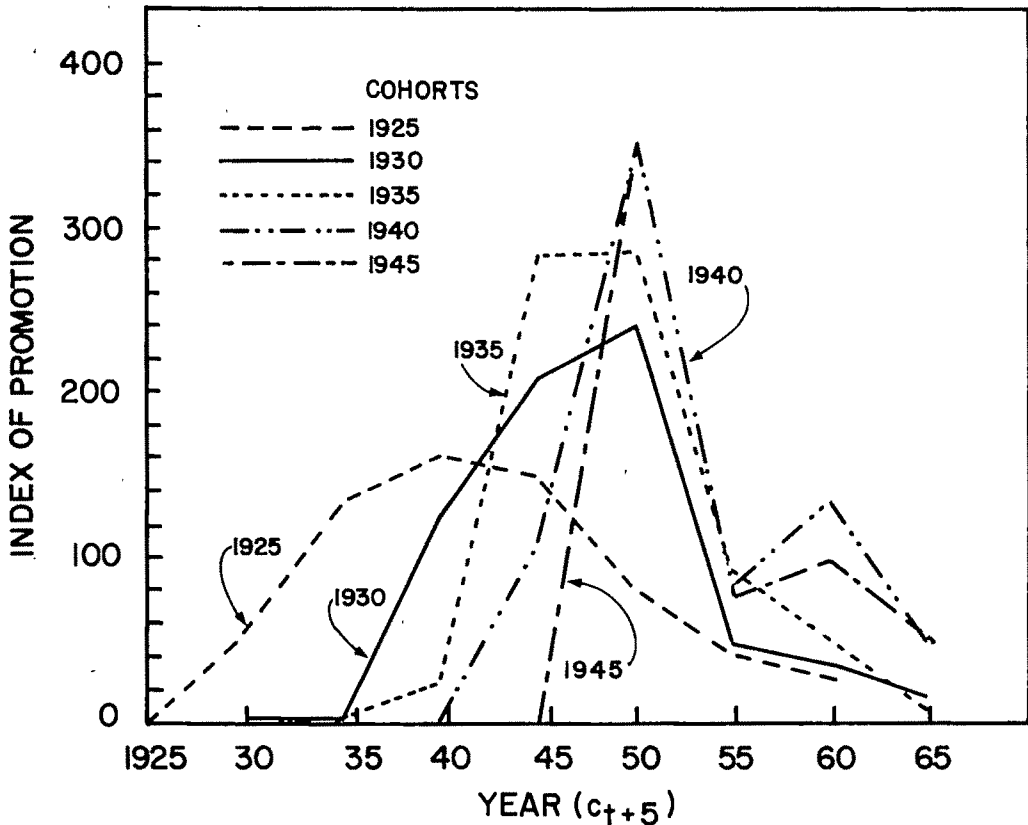


Figure 1. Index of Promotion for Five Year Cohorts of Foreign Service Officers, 1925-1965

this particular response was chosen, one must examine cohorts of FSOs more closely as they pass through this system. An examination of two possible alternative explanations of these results eliminates them as major contenders for interpreting the results.

First, although vacancy patterns are not as critical for understanding career mobility in the Foreign Service as in the clerical careers studied by White (1970), attrition rates were examined nevertheless to see if there were pronounced changes which might explain the intercohort differences observed above. Increased possibilities for promotion during the war years led to lower age specific attrition rates as officers stayed in the Service longer, slightly diminished the size of the vacancy pool, and hindered the continuation of the 1940-1950 promotion rates into the

1950s.¹³ Figure 3, which gives these seniority-specific attrition rates, reveals surprising consistency in the patterns over time for each of the cohorts. Relatively little interaction between promotion rates (Figure 2) and attrition (Figure 3) is revealed in the 1930, 1935, 1940, and even 1945 cohorts. The general pattern is in the direction anticipated, but the relative

¹³ The institution of the two new supergrades, Career Minister and Career Ambassador, in 1947, may have contributed to the rise in promotion rates through the creation of longer vacancy chains. By the early 1960s, the lack of upper level vacancies was identified as a problem of "not enough jobs for talented, experienced senior officers" (Harr, 1969). Although there is not a direct correspondence between job positions and openings at different career grades, it is clear that a marked expansion of opportunities for promotion has not been part of the Foreign Service response to the changing international environment. The dominant modes of response included some functional specialization, expansion of language skills, and the broadening of the recruitment base.

while generally useful, suggest this particular form of adaptive response.

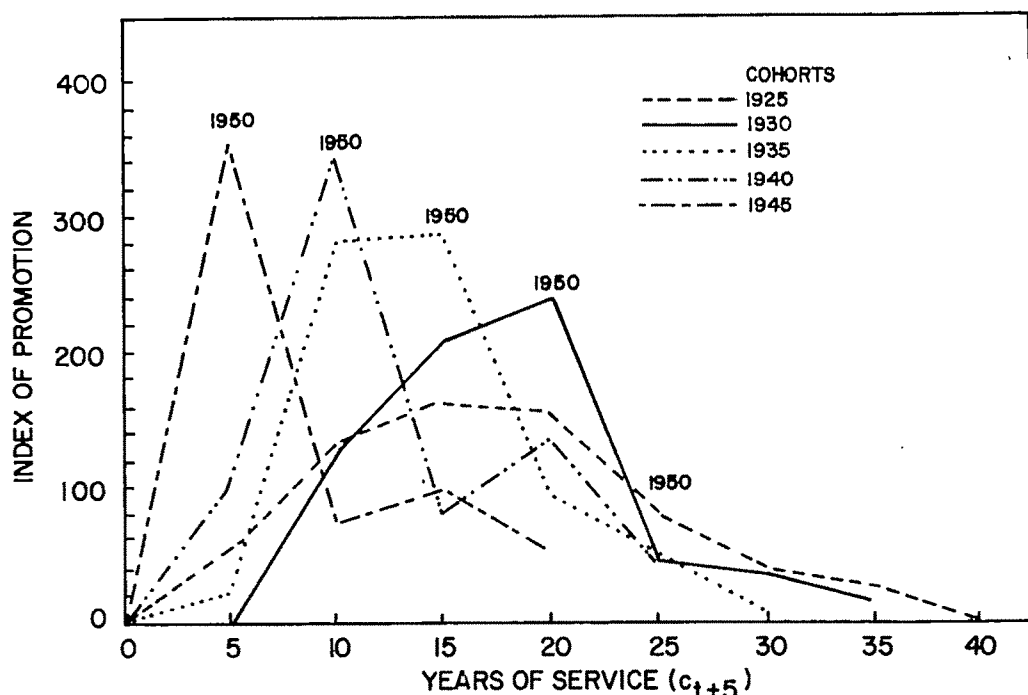


Figure 2. Index of Promotion for Five Year Cohorts of Foreign Service Officers (1925–1945): Showing Changes in Promotion Rates of Successive Cohorts

magnitudes of change are not directly proportional to fluctuations in promotion rates.

Second, intercohort differences could be a byproduct of rule changes which accelerated the rate at which younger FSOs reach managerial or ambassadorial rank. The data presented in Table 1, which shows the average promotion indices for cohorts after twenty years of service, do not support this interpretation. Although there is considerable internal variation in promotion indices, there is substantial consistency in average promotion rates (hence, average rank) for all cohorts completing twenty years of service. The existence of the intracohort variation suggests that decisions about access to senior leadership positions are highly buffered from external interferences (Hickson et al., 1971; Hinings et al., 1974).¹⁴ Net flows of people to senior positions because they are defined by the career system, lie

within the scope of administrative control (Stinchcombe et al., 1968). Rather than grow through the staffing of new foreign affairs positions, assume new functions, or open the career system to substantial and continued lateral entry, the Foreign Service Administration maintained the existing size, composition, and rank structure for the Service, which favored the preservation of both the traditional diplomatic functions and the elite character of the generalist Officer Corps. In general, a reequilibration of promotion rates could be expected to insure continuity in succession, thereby minimizing the possibility of major shifts in organizational policy as a result of new and inadequately socialized personnel (McNeil and Thompson, 1971).

Within the constraints of a closed-career system, there are still options for minor adjustments in response to external mandates. An examination of Foreign Service Officers who completed twenty years of service reveals several of important adaptations within the general framework of a traditional Officer Corps already outlined. First, no officers were

¹⁴ Such measures provide independent evidence of the central technology of the Foreign Service. According to Thompson's (1967) theoretical scheme, these are expected to be highly buffered from environmental fluctuations.

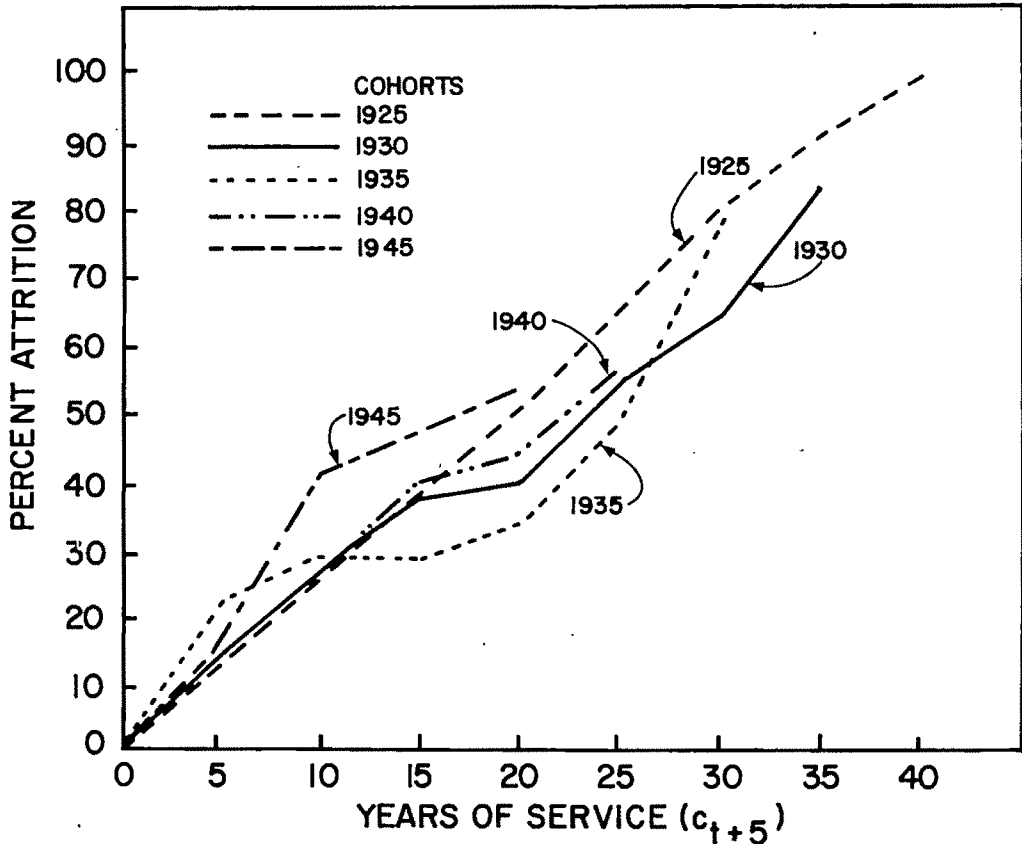


Figure 3. Cumulative Attrition Rates for Five-Year Cohorts of Foreign Service Officers

promoted to Career Minister or Career Ambassador positions with less than twenty years of service. This, as we discussed earlier, is consistent with the conception of the traditional Foreign Service career.¹⁵ At the same time, there has been a slight trend toward the promotion of younger officers to these senior (Career Minister and Career Ambassador) positions. Second, with each successive cohort, a greater percentage of those entering the Service go on to become Career Ministers or Career Ambassadors. This modest change is coupled with a slight decrease in the proportion of members of each cohort rising to these senior positions later in their careers. Thus, de-

spite retention of the closed cohort-based career system and the buffering of access to senior positions, the Foreign Service Administration has been sensitive to pressures for change in the leadership of the Officer Corps.

SUMMARY AND CONCLUSIONS

The longitudinal study of organizational reproduction and change in the American Foreign Service reveals the analytic contribution of the cohort approach to the study of complex organizations. Inter-cohort differences in composition, experience and transformation over time provide important structural variables for the analysis of organizational change.

In developing the cohort perspective, this paper has explored the utility of three propositions generated from the existing literature on cohort demography. Simultaneously, the perspective makes it possi-

¹⁵ There was a change from a traditional pyramidal to a distribution of ranks with a "bulge in the middle" characteristic of bureaucracies developing greater specialization. This has led to strong sentiment for a return to a smaller, more elite, streamlined Officer Corps (Harr, 1969:264-6).

ble to organize a number of discrete observations about developments in the modern Foreign Service.

The propositions stated at the outset of this paper were explored in a review of previous studies of the Foreign Service and in the analysis of data on the outcome of promotion practices in the Officer Corps.

The first proposition, on the lag time consequences of cohort organization, is difficult to test definitively without access to greater data resources. Nevertheless, it has been possible (1) plausibly to establish the existence of a time lag between environmental change and the adaptation of the Foreign Service to changes in its environment, and (2) to identify the contribution of cohort composition to delayed adaptation. The variety of ancillary structures (e.g., the Foreign Service Auxiliary) and temporary adaptations in the career system (e.g., lateral entry) further attest to both the need for adaptation and the difficulty of making those changes within the context of the closed-career system.

The second proposition, stating that any particular cohort career system seriously constrains the distribution of organizational rewards to individuals, is also used to guide the analysis of organizational reproduction in the American Foreign Service. A variety of internal conflicts and a number of external evaluations of the department identified blocked career mobility as a recurrent issue during the late 1960s. At the same time, however, commitment to the particular career system made it difficult for the Administration of the Foreign Service to cope with the consequences of past recruitment and promotion commitments.

In examining the third proposition, I proposed that the particular cohort career form and preferred demographic distribution became an important tool in the organizational political economy. While we cannot address the full details of this proposition, we can bring to bear the following information that supports its conclusions. First, the data in Figure 1 reveal that, though there is variation between cohorts in promotion rates, these differences are not sufficient to transform dramatically the personnel structure of the

Foreign Service. Second, the data in Figure 2 reveal that even the time-in-service career pattern of newer cohorts conforms closely to the general pattern of older cohorts. This is not a situation calculated to transform the personnel structure. If the career structure were to be transformed dramatically by the recruitment of officers with new skills and commitments, then these figures should show that newer cohorts are promoted much more rapidly (Table 1 reveals this not to be the case), that they move higher in the organization per unit time (they do not), and that at the end of twenty years they would have had higher rather than lower average promotion indexes (see Table 1). The analysis of promotion during and immediately after World War II revealed that maintenance of the career system delayed the rise of younger officers to senior policy positions and helped to maintain an elitist Officer Corps focused on traditional diplomatic functions.

Cohorts, reflecting different degrees of seniority in the organization and unique transformational processes, are thus a central unit for the study of social change, and reflect a broader need for the examination of organizational demography (Hannan and Freeman, 1977). As the bulk of recent research anticipates, the reproductive patterns in social systems reflect the political economy of the particular organization and the structurally determined differences in ability to control strategic consequences. These processes may be critical for understanding the internal politics of organizations as cohorts acquire political interests which are different and/or mutually contradictory. Technological changes may make members of older cohorts obsolete as senior organizational officials rationalize operations or redefine the organizational mission. Alternatively, overall organization policy may reject new developments. Then, new competitors or new organizational units will be created to utilize the new technologies, skills, and associated needs of younger cohorts. The Foreign Service, and those senior officers who administer it, reasserted control over the processes of organizational regeneration during the period under study,

through the preservation of the closed-career service. Adaptive responses were confined to other areas of Foreign Service policy. The implications of this pattern of control for the analysis of adaptation in complex organizations and for the development and execution of foreign policy should be the subject of further investigation.

In summary, the cohort perspective provides us with specific guidelines for inquiry while at the same time it offers a general middle-range theoretical framework for understanding important constraints associated with organizational commitment to a particular career structure. Professionalization and elaborated work rules may make personnel transfers at middle and upper levels of organizations increasingly difficult. As a consequence, major shifts in the orientation of large scale organizations are dependent more on recruitment and socialization at lower levels of the organization on the one hand,¹⁶ and on those processes which in-

fluence cohort transformation—age-specific attrition rates, vacancy patterns, career experiences, and promotion—on the other.

Cohorts provide a viable analytic unit for understanding certain aspects of complex organizations. Pressures for regeneration are frequently associated with newer cohorts acting as the bearers of technological innovations and new values. The responsiveness of organizations is contingent on the relative size and structure of cohorts and their impact on mobility into positions of organizational control. Finally, resistance to change is associated frequently with older cohorts whose skills and interests may become obsolete in a changing world. How pervasive or important are these cohort phenomena? There is a growing body of evidence to suggest that the perspective of organizational demography will contribute much to any analysis of organizational change.

are controlled by those whose commitments and skills may be jeopardized by these very changes.

¹⁶ Both recruitment and socialization frequently

APPENDIX

PERCENTAGE DISTRIBUTION OF THE FOREIGN SERVICE OFFICERS CORPS, BY RANK, FOR SELECTED YEARS*

Personal Rank/Grade	1926	1931	1936	1941	1945 ^b	1951	1956	1961	1966
Career Ambassador	— ^c	— ^c	— ^c	— ^c	— ^c	— ^c	1	1	1
Career Minister	— ^c	— ^c	— ^c	— ^c	— ^c	1	2	2	2
FSO-1	4	5	5	5	6	4	5	6	8
FSO-2	4	5	5	5	6	10	11	12	12
FSO-3	8	7	5	7	10	16	20	16	19
FSO-4	9	9	9	10	10	19	30	19	18
FSO-5	10	10	9	10	11	22	21	17	13
FSO-6	12	13	14	12	15	28	11	8	14
FSO-7	14	13	12	16	21	— ^c	— ^c	10	10
FSO-8	15	10	12	11	9	— ^c	— ^c	9	5
Unclassified	24	29	29	24	15	— ^c	— ^c	— ^c	— ^c
Total	653	703	683	857	790	1363	2515	3708	3507

* These figures were compiled for the entire population of the Officer Corps, not just the cohort sample used elsewhere as a basis for analysis. Percentages are rounded.

^b Data are not available by rank for 1946. Comparable figures for 1945 are given instead.

^c During this year this rank did not exist. Changes in the ranking system for the Officer Corps have occurred at several times. An account of one such change is given by Steiner (1958:29),

In 1956, the Foreign Service Act of 1946 was amended to increase the number of classes from seven to ten by adding two classes at the bottom and by splitting the enlarged FSO-4 class in two. The change was made to end the undue congestion at the middle levels and to give more flexibility and balance to the total structure.

No officers in the sampled cohorts were changed in rank as a result of this or other modifications in the ranking system.

REFERENCES

- Barnes, William, and John Heath Morgan
1961 "The Foreign Service of the United States: origins, development, and functions." Historical Office, Bureau of Public Affairs. Washington, D.C.: U.S. Government Printing Office.
- Benson, J. Kenneth
1975 "The interorganizational network as political economy." *Administrative Science Quarterly* 20:229-49.
- Brookings Institution
1951 *The Administration of Foreign Affairs and Overseas Operations*. Washington, D.C.: U.S. Government Printing Office.
- Commission on the Organization of Government for the Conduct of Foreign Policy
1975 Report. Washington, D.C.: U.S. Government Printing Office.
- Committee on Foreign Affairs Personnel
1962 *Personnel for the New Diplomacy*. New York: Carnegie Endowment for International Peace.
- Dahrendorf, Ralf
1968 "On the origin of inequality among men." Pp. 151-78 in Ralf Dahrendorf (ed.), *Essays in the Theory of Society*. Stanford: Stanford University Press.
- Deutsch, Karl
1963 *Nerves of Government: Models of Political Communication and Control*. New York: Free Press.
- Downs, Anthony
1967 *Inside Bureaucracy*. Boston: Little, Brown.
- Feinstein, Harry
1969 *A Study of the Pattern of Promotions from Class FS06 to Class FS03 of Foreign Service Officers*. Unpublished General Studies Paper, George Washington University.
- Hall, Douglas T., and Benjamin Schneider
1973 *Organizational Climates and Careers: The Work Lives of Priests*. New York: Seminar Press.
- Hannan, Michael T., and John Freeman
1977 "The population ecology of organizations." *American Journal of Sociology* 82:929-64.
- Harr, John E.
1965a *The Anatomy of the Foreign Service*. New York: Taplinger.
1965b *The Development of Careers in the Foreign Service*. New York: Taplinger.
1969 *The Professional Diplomat and the New Diplomacy*. Princeton: Princeton University Press.
- Hass, J. Eugene, and Thomas E. Drabek
1973 *Complex Organizations: A Sociological Perspective*. New York: Macmillan.
- Hernes, Gudmund
1976 "Structural change in social processes." *American Journal of Sociology* 82:513-47.
- Heydebrand, Wolf V.
1975 "Organizational contradictions in public bureaucracies: toward a Marxian theory of organization." Paper presented at the annual ASA meeting, San Francisco.
- Hickson, D. J., C. R. Hinings, C. A. Lee, R. E. Schneck, and J. M. Pennings
1971 "A strategic contingencies theory of intra-organizational power." *Administrative Science Quarterly* 16:216-29.
- Hinings, C. R., D. J. Hickson, J. M. Pennings, and R. E. Schneck
1974 "Structural conditions of intra-organizational power." *Administrative Science Quarterly* 19:22-44.
- Hoover Commission
1949 *Foreign Affairs: A Report to the Congress by the Commission on Organization of the Executive Branch of the Government*. Washington, D.C.: U.S. Government Printing Office.
1955 *Task Force Report on Overseas Economic Operations*. Washington, D.C.: U.S. Government Printing Office.
- Inglehart, Ronald
1971 "The silent revolution in Europe: intergenerational change in post industrial society." *American Political Science Review* 65: 991-1017.
- Janowitz, Morris
1960 *The Professional Soldier*. Glencoe: Free Press.
- JFSOC
1967 *Junior Foreign Service Officer Club: Survey Results*. Mimeo. Washington, D.C.
- Jones, Arthur G.
1965 *The Evolution of Personnel Systems for U.S. Foreign Affairs: A History of Reform Efforts*. New York: Taplinger.
- Kanter, Arnold
1976 "Managerial careers of Air Force generals: a test of the Janowitz convergence hypothesis." *Journal of Political and Military Sociology* 4:121-33.
- Kaufman, Herbert
1960 *The Forest Ranger*. Baltimore: Johns Hopkins Press.
- Lipset, Seymour Martin
1959 *Agrarian Socialism: The Cooperative Commonwealth Federation in Saskatchewan*. Berkeley: University of California Press.
1960 *Political Man: The Social Bases of Politics*. Garden City: Doubleday.
- McNeil, Kenneth, and James D. Thompson
1971 "The regeneration of social organizations." *American Sociological Review* 36:624-37.
- Macomber, William B.
1970 *Diplomacy for the '70's: A Program of Management Reform for the Department of State*. Washington, D.C.: U.S. Government Printing Office.
- Mannheim, Karl
1952 "The problem of generations." Pp. 276-322 in P. Kecskemeti (ed.), *Essays on the Sociology of Knowledge*. New York: Oxford University Press.
- Martin, Graham
1968 *Toward a Modern Diplomacy: A Report to the American Foreign Service Association*. Washington, D.C.: American Foreign Service Association.

- Michels, Robert
1958 *Political Parties*. Glencoe: Free Press.
- Moskos, Charles C., Jr.
1970 *The American Enlisted Man: The Rank and File in Today's Military*. New York: Russell Sage.
- Nagle, John D.
1975 "A new look at the Soviet elite: a generational model of the Soviet system." *Journal of Political and Military Sociology* 3:1-13.
- Perrow, Charles
1972 *Complex Organizations: A Critical Essay*. Glenview: Scott, Foresman.
- President's Task Force on Government Organization
1967 *Organization for Foreign Affairs*. Washington, D.C.: U.S. Government Printing Office.
- Reed, Theodore L.
1972 *Structure and Process in Organizational Careers: A Computer Simulation of Promotion in the American Foreign Service*. Ph.D. dissertation, Department of Sociology, University of Michigan.
- Riley, Matilda White
1973 "Aging and cohort succession: interpretations and misinterpretations." *Public Opinion Quarterly* 73:35-47.
- Ryder, Norman B.
1965 "The cohort in the study of social change." *American Sociological Review* 30:843-61.
- Scott, Andrew M.
1969 "The Department of State: formal organizational and informal culture." *International Studies Quarterly* 13:1-18.
- Segal, David R., and Daniel H. Willick
1968 "The reinforcement of traditional career patterns in agencies under stress." *Public Administration Review* 28:30-8.
- Steiner, Zara S.
1958 *The State Department and the Foreign Service: The Wriston Report Four Years Later*. Princeton: Center of International Studies, Princeton University.
1961 *Present Problems of the Foreign Service*. Princeton: Center of International Studies, Princeton University.
- Stinchcombe, Arthur L.
1965 "Social structure and organizations." Pp. 142-93 in James G. March (ed.), *Handbook of Organizations*. Chicago: Rand McNally.
- Stinchcombe, Arthur L., Mary Sexton McDill, and Dollie R. Walker
1968 "Demography of organizations." *American Journal of Sociology* 74:221-9.
- Tepperman, Lorne
1975 "Demographic aspects of career mobility." *Canadian Review of Sociology and Anthropology* 12:136-77.
- Thompson, James D.
1967 *Organizations in Action*. New York: McGraw-Hill.
- Warwick, Donald P. in collaboration with Marvin Meade and Theodore L. Reed
1975 *A Theory of Public Bureaucracy: Politics, Personality and Organization in the State Department*. Cambridge, Ma.: Harvard University Press.
- Weber, Max
1947 *The Theory of Social and Economic Organization*. New York: Free Press.
- White, Harrison C.
1970 *Chains of Opportunity: System Models of Mobility in Organizations*. Cambridge, Ma.: Harvard University Press.
- Wright, Eric Olin, and Luca Perrone
1977 "Marxist class categories of income inequality." *American Sociological Review* 42:32-55.
- Wriston Committee (Secretary of State's Public Committee on Personnel)
1954 *Toward a Stronger Foreign Service*. Washington, D.C.: Department of State Publication, No. 5458.
- Zald, Mayer
1970a *Organizational Change: The Political Economy of the YMCA*. Chicago: University of Chicago Press.
1970b "Political economy: a framework for comparative analysis." Pp. 221-61 in Mayer Zald (ed.), *Power in Organizations*. Nashville: Vanderbilt University.
- Zeitlin, Maurice
1970 *Revolutionary Politics and the Cuban Working Class*. New York: Harper and Row.

STYLE AS SOCIAL PROCESS*

BARBARA ROSENBLUM

Stanford University

American Sociological Review 1978, Vol. 43 (June):422-438

This comparative study of news, advertising and fine arts photographers argues that stylistic conventions in each of three categories of photography are partly determined by basic structural processes which, in turn, shape photographers' control over technical and aesthetic aspects of the photographic process. This study provides a meaningful approach to the analysis of material culture, an area which has been investigated from points of view other than sociology.

In recent years, interest in the sociology of art has grown tremendously.¹ However, a review of empirical studies within this substantive area reveals a concentration of efforts on a few basic questions such as the socialization of art students (Fields, 1976; Frederico, 1974; Griff, 1964; Hecht, 1971; Smades, 1976; Schuyler, 1976) or the impact of technology on various media (Bennett, 1972; Kealy, 1974; Hennessey, 1975).

In contrast, studies which take into account features of the art object itself are not usually done in the sociology of art.² There are many difficulties associated with such an endeavor. It requires that the sociologist become thoroughly conversant in the lexicon of the medium studied and also become somewhat of an expert in stylistic analysis. Although this is not easy, it is essential if we are to build a systematic approach to the sociological study of style and to avoid empirical pitfalls that are potentially detrimental to theorizing. For instance, recently Peterson and Berger (1975) advanced the ex-

planation that the stylistic homogeneity of popular music is a function of market concentration. But because their measures of homogeneity are indirect (for instance, the chart position of a recording) and have nothing to do with the music itself, the plausibility of their assertions about music styles is difficult to assess. The investigation of the frequency, structural articulation and location of flatted thirds or perfect fourths (respectively, features of American blues and of Eastern European folk songs) is necessary in order to make assertions about music styles. We must pay attention to the content and structure of art products if we are to understand how social processes shape material culture (Kroeber, 1948; Wolff, 1975).³

In this study, I look at stylistic features of photographs and try to account for their patterned regularities by focusing on aspects of the social organization of the production of photographs in three settings: news, advertising and fine arts. In addition to looking at stylistic features of the photographs, I also look at the range of variation there is within each category of photography. I propose that, in addition to other explanations of style, style is also a function of the structural characteristics and constraints associated with typical situations in which photographs are made. This approach is significantly different from explanations which see style as a product of large scale sociohistorical

* I would like to thank the following people for their critical comments on earlier drafts of this paper: H. Stith Bennett, Francesca Cancian, Rachel Kahn-Hut, John Meyer, Chandra Mukerji, and Morris Zelditch, Jr.

¹ There is now an annual sociology of arts conference which has been organized independently of any professional sociological association.

² This assertion should be qualified by the comment that I am limiting my discussion to American empirical sociology of art. European social theorists, such as Lukacs and Habermas have in fact paid attention to features of art products (see Burns and Burns, 1973).

³ A convincing argument for the need for the sociology of art to incorporate analyses of art products has been put forth recently by Wolff (1975).

forces such as the commodification of art products in capitalist societies (see Baxandall, 1972; Burns and Burns, 1973; and Fischer, 1963; for explications of a Marxist approach). My approach is also significantly different from the recent work of Becker (1974) and Burns (1972). Although neither of them focuses specifically on style, their emphasis on the concept of convention has implications for the study of style. Rooted in a symbolic interactionist framework, each argues that conventions or shared agreements is the principal independent variable that accounts for recurrent patterned activity and, by extension of their argument, the objective products of those activities. Their work, however, stops short of analyzing what, if any, larger social processes shape the development of conventions.⁴ I, too, am interested in the concept of convention but I treat it differently from each of them. Where they treat it as an independent variable, I treat convention as an intervening variable, sandwiched between the organization of production and the characteristics of the final outcome. Like Becker (1974) and Burns (1972), I assume that there are shared meanings and standardized ways of doing things. Like Kroeber (1948:137), I assume that standardized ways of doing things produce a given style. But, unlike them, I believe that shared agreements are not independent of more basic features and structures of social life. Rather, shared agreements arise from basic social structural arrangements, on the one hand, and shape them, on the other hand. In other words, I am seeking a basis for convention which takes the simultaneous reciprocal influence of social structure and institutionalized understandings into account. Consequently, I treat convention as a function of social organization. One implication is that conventions vary with contexts of production. Furthermore, each context generates different styles of photography. Put another way, one should understand the social ar-

rangements in which pictures are made in order to understand why they look as they do.

To understand how conventions are shaped by economic, technical, political and social constraints, I will focus on what photographers actually do in each setting and delineate the basic social processes that shape their control over making pictures. Since convention is, in one sense, a product of human labor, it is essential to explore the nature of picture making by looking at what photographers can and cannot do in each setting. Thus, by comparing how photographers make pictures in each of three settings, I show that stylistic differences between news, advertising and fine arts photographs may be partially attributed to the organizational differences between those three settings.

To show how conventions and understandings feedback to shape structural and organizational arrangements, I will look at institutional features of journalism, advertising and fine arts. This study analyzes styles in photography by looking at both organizational and institutional influences simultaneously. This approach, then, attempts to modify theories which tend to overemphasize the cultural autonomy of style, its independence from social structure and the "out-there-ness" of socially unlocated shared meanings.

RESEARCH METHODS

The bulk of the data was collected in New York between 1971 and 1972. Prior to that, I did pilot interviews in Chicago during 1970-71. Subsequent to the research period in New York, I collected additional data in both San Francisco and New York to update my findings.

The principal method of data collection was participant observation. Especially important in this research was the development of long-term relationships with key informants who were invaluable resources all through this research. The materials gathered from participating and observing provided the basis for semifocused interviews I administered to a sample of photographers whom I did not previously observe. That sample of formally interviewed photographers was generated

⁴ The philosopher Benedetto Croce (1968:31) puts the problem this way, "But if there are to be conventions, something must exist which is no convention but is itself the author of convention."

by referrals from those photographers I did observe at work. Although I attempted to construct a random sample, various sampling difficulties prevented me from doing so. For instance, I was often denied an interview when I did not have a referral. As a result, I used snowball sampling and tried to control for error in a variety of ways (see Rosenblum, 1978, for details).

I also conducted formal interviews with people with whom photographers frequently come in contact in each of the three settings. In news, I interviewed editors, reporters, darkroom personnel, and typographers. In advertising, I spoke with models, art directors, photographers' representatives and others. For fine arts photography, I interviewed museum curators, gallery directors and art school personnel. These additional interviews proved invaluable for they provided me with details about work organization and role relationships that the photographers neglected to tell me. Finally, because I needed to know photographic technique in order to ask photographers why they photographed something one way and not another, I took photography courses.

STYLES IN PHOTOGRAPHY

Style consists of particular mannerisms or conventions that are frequently associated together. Because of the frequency of association, we treat the entire combination as an identifiable whole which is generally consistent and stable over time and place. One sociologically useful way to think about style is that it consists of predictable combinations of features.

Styles in photography are empirically unproblematic. Most people, when asked to distinguish types of photographs from one another, can do so. Most people can recognize a news picture and distinguish it from a picture in an ad and, furthermore, distinguish each of those from a picture that hangs in a museum.⁵ Although there is some borrowing and exchange between

these categories, I am assuming the distinctiveness and independence of these three styles of photography.

Each of these styles may be analyzed with respect to its component features. Even for one untrained in stylistic analysis, it is possible to use one's eyes and powers of observation to identify these component features and describe them.

News pictures are generally recognizable by their specific range of content, depicting routine events such as baseball games, press conferences, and urban violence. The people in news pictures are typically in focus, clearly identifiable and almost always are photographed in the situation which, itself, has been defined as newsworthy. The composition of the picture is perhaps its special signature. The center of the composition is nearly always occupied by the key newsworthy figure. People and objects are usually arranged so that the informational content of the photograph constitutes its primary impact. Frontal positioning, dominant verticals, linear-planar spatial construction and unvaried surface pattern are some of the main aesthetic ingredients of news photography. Nuance, subtlety, shading and mood are typically absent from news pictures. Also absent are elements which would vitiate its informational impact. There are, of course, occasional exceptions. In fact, some of the most famous news pictures are out of focus, e.g., the Pulitzer prize winning photograph by Eddie Adams which depicted a South Vietnamese captain shooting a suspected Vietcong terrorist. By and large, such photographs are exceptions, and typical news pictures partake of the features described here.

Advertising photography, in contrast, depicts extraordinary people or objects in extraordinary situations presented in visually extraordinary ways. A typical example of such a photograph is a picture of a brand new, highly polished automobile placed in the middle of a huge grassy meadow where a model, dressed in a flowing crepe, floorlength evening dress, stands near the car, looks up towards the sky and smiles at us. We find in advertising photography may unconventional ges-

⁵ In one of my classes, I asked students to sort 20 photographs with respect to whether they thought it was a news, advertising or fine arts photograph. Agreement was 97%.

tures, postures and visual effects. Not all people or objects or situations in advertising are extraordinary; sometimes ordinariness oozes from the image. What we do find in advertising photography are extraordinary visual situations which play with our senses of scale, depth and distance. The hallmark of advertising photography is the treatment of three dimensional space: short foregrounds, complex middlegrounds and deadstop, horizonless backgrounds. Although space limitations here prevent a complete discussion of the technical construction of advertising space (see Rosenblum, 1978, for details), it is clear that unnaturalness is the principal aesthetic component of advertising photography. Evidence for such can not only be seen in spatial construction but also in the hypertactile and atmospheric effects of the genre.

Fine arts photography is perhaps the most elusive of all types of photography and the most difficult to describe with a few words. The range of imagery included under the heading "fine arts photography" is much greater than for news and advertising. Consequently, problems are created for the stylistic analysis of fine arts photographs because of the elasticity of the boundaries. In a sense, one could view fine arts photography as a large residual category which subsumes a good many types of unclassifiable photographs. However, one would be incorrect to make such an inference because, while the boundaries are elastic, the boundaries do not admit any and all combinations of stylistic conventions. Rather, a selective process is at work.

One of the key features of fine arts photography is the self-conscious representation of space, meaning and light. In other words, the representation of space, meaning and light becomes the content of the image, no matter what the subject matter. Space is especially important and fine arts photography is permeated by what is called a *space* sensibility (Read, 1965:64) or abstract space (Gombrich, 1960) rather than the *place* sensibility that is found in both news and advertising. Another characteristic of fine arts photography is the stamp of the personal vision of the photographer which can often be an indi-

vidualized, sometimes introverted view of the world which does not rely on a conventionalized set of symbols for its meaning (see Langer, 1942, for a discussion of discursive and nondiscursive symbolism).

In addition to these features associated with each style, there is another dimension that is important to consider. This dimension concerns the range of variation within each category. Even for persons unfamiliar with the history of photography, a comparison of news, advertising and fine arts photography quickly shows that each may be ranked with respect to its degree of homogeneity. Images allowable under the rubric of photojournalism are more narrowly circumscribed than those which fall under advertising. As we already saw, fine arts photography is distinguished by a pluralism of aesthetics and so may be considered more fluid than the other two. Let us now examine the social organization of the three settings in which these types of photographs are made.

NEWS

Newspaper photographers,⁶ usually when they come into the office to begin their day, are given their assignments by a photo assignment editor. The beginning of the day varies with the photographer's shift and the particular newspaper—whether it is a morning or evening paper, the number of editions it puts out. A photographer is typically assigned between two and four events to cover in a single eight-hour day. These are prescheduled, routine events, which may range from a portrait of a recently promoted bank vice-president (for the financial pages) to a weather picture (a possible page one). The photographer's day, then, consists minimally of all the things he⁷ must do to complete his assignments: travel to the scenes of the events, assess

⁶ Photographers who are employed by newspapers are sometimes called photojournalists and their photographs are often called reportage. But because these terms invoke traditions of style and may be confusing, I shall simply call them newspaper photographers here.

⁷ I use the term "he" because newspaper photography is a male occupation. Only 2 of the 68 New York photographers were women in 1973.

newsworthy aspects of each event, select and record those visual images which especially capture the action and convey the meaning of the event, and return to the office with rolls of film that, when processed and printed, can be useful to the picture editors who select the photographs to accompany stories. Each of these activities is a learned, complex activity, the knowledge for which is acquired during the photographer's apprenticeship period.

To do his job, the photographer is issued company equipment, consisting of two single-lens reflex camera bodies and three lenses, one normal lens (generally 50 mm but sometimes defined as 35 mm), one wide-angle lens (the 24 mm lens was typical for my sample) and one telephoto lens. Most newspapers use standard equipment, although the type of telephoto lens varied from newspaper to newspaper. Special lenses, such as a catadioptric lens, are theoretically available to any staff photographer from a central supply cabinet, but these are often under special lock-and-key and requisitions must be approved before the equipment can be used. By the time the bureaucratic paperwork has been accomplished, the event may have already passed. Thus, for the most part, newspaper photographers all use the same equipment.

In addition to equipment, film is standardized as well. Film rated as ASA 400 was the only film I ever saw newspaper photographers use. Its flexibility in a variety of lighting situations is the characteristic that recommends its use. By now, however, its use is taken for granted since all the automated film processing equipment is coupled to match that film. Automated film processing equipment is capable of transforming an exposed roll of film into a dry, uncut strip of negatives in six minutes from start to finish. A photograph can make an edition if its negative arrives only an hour or less before the presses begin.

The photographer's lack of control over the developing and printing processes is directly attributable to automation. The photographer cannot select his own favorite chemicals nor determine their temperature nor exercise decision making as in hand processing. In photography, each

technical decision is also an aesthetic decision and contributes to the total outcome, that is, to what the picture looks like. Thus, in newspaper photography, pictures look very much alike in terms of technical qualities because they have all been processed similarly.

The proof sheets when printed are passed to a photo editor who reviews them and selects from among all the available images those which will be enlarged for possible reproduction. The darkroom staff in turn enlarges and prints the editor's selections in accordance with their technical criteria for good reproduction. These technical criteria are matched with—and perhaps emerged from—the specifications and characteristics of the reproduction process.

When the enlarged picture is returned to the photo editor, he may study it, crop it, cut out some background details or send it back to the darkroom for a more detailed enlargement. In short, the photo editor selects which photographs will be used and how much of the image will be used and how much column space will be allotted to key photographs. He changes the picture to make it more suitable, more graphically and informationally relevant to the story it will accompany. From there the photograph itself passes through a variety of people, from retouchers to typographers, each of whom makes technical decisions about how the picture will look in print. Even at the very last stage, typographers can exercise some control over the contrast of the image, i.e., the relationships between the blacks, whites and some intermediate tones.

From this brief description, it is possible to discern several structural and organizational characteristics in the world of newspaper photography. In the first place, the entire photographic process is dominated by technology. I have already pointed out that in photography every technical decision affects the final outcome, what the photograph will look like. Automation and mechanization alone are responsible for the many regularities we see in newspaper photographs. In comparison, advertising and fine arts photographers determine the outcomes by retaining their technical control over the

picture-making process. In news photography, the sequences of processing and the treatment given each stage are predetermined and standardized, whereas in advertising and fine arts, the identical stages are not mechanized. Undoubtedly, multistage mechanization arises from the pressure of edition deadlines and this, of course, affects the newspaper photographer and his work.

Newspaper photography is also characterized by the division of labor, the usual companion of automation. Each of the steps in the picture-making process, which was traditionally the photographer's sole province and area of expertise, is allocated to a variety of persons, each of whom is an expert in a small piece of knowledge. Thus, in news, the picture-making process from beginning to end is a highly complex, coordinated series of activities and communications that spreads out decision making over a considerable number of people. The way the picture finally looks, then, is the result of a kind of assembly line production in which individual persons make discrete decisions about pictures. Furthermore, these decisions are typically not made in consultation or in concert with one another.

We do know from previous studies of bureaucracies (see Crozier, 1964; Meyer, 1971) that bureaucratic organizations generate their own classification systems. As a result, shorthand categories come to stand for ways of doing things. On the organizational level alone, a bureaucratically organized newspaper's reliance on its own category system tends to homogenize photographic imagery simply through rejection of images which do not fit the existing category system. But each newspaper's system of categories is shaped by the effects of the pan-organizational system of institutional journalism. In fact, the organization of the division of labor within each newspaper is partly shaped by the institution of journalism. It is here that the dialectic between social structure and institution is perhaps most evident.

Journalism is characterized by a national system of categories that are institutionally pervasive. Examples of such

categories are disasters, elections, celebrities, heroic stories, horror stories, medical breakthroughs and so on. Thus, while the dramatis personae may change over time as public officials, baseball players and murderers come and go, the basic news scenarios remain the same. The institutionally pervasive system of journalistic classification may be seen as a highly elaborated code (Bernstein, 1971) that permits the differentiation of all human activity into subunits. On the intraorganizational level, there is a correspondence between functionally specific departments and roles and many of the journalistic categories. I do not mean to imply that there is total isomorphism between journalistic categories and the social division of labor, but there is a strong relationship which varies with the size and circulation of each newspaper.

Each newspaper, then, is dependent upon the institutional system of journalistic categories and must intermesh its own local and idiosyncratic categories with those that are institutionally prevalent. The combined effects of an institutionally pervasive codification of the world, in addition to the bureaucratic strain towards calculability and predictability, yields a picture of a news organization that relies heavily on shared agreements, especially shared agreements which concern the matching of real life events with news categories (see Merton, 1957; Schutz, 1964; 1971; and Berger and Luckmann, 1966, for explications of classification systems as reality-maintaining frameworks and as decision-making frameworks). Thus, the social organization of institutional journalism generates its own tendency towards redundancy through the elimination of many uncodable or insignificantly evaluated real life events (see Tuchman, 1973; and Molotch and Lester, 1974, for a social construction of news approach).

The system I have just described affects how photographers take pictures and how photographs are selected. Both should be discussed together. The one major purpose of a news photograph is to illustrate aspects of the content of a story. But remember that every specific story is treated as a concrete instance of a larger

abstract category. The abstract category is an ideal type of all similar stories. And its properties are typified as well. Because the abstract category remains the same, even though the content of the real life stories vary somewhat, photographs that illustrate properties of the ideal type are repeatedly selected for illustration. This is an organizational and institutional dynamic that contributes towards the homogeneity of imagery we find in newspapers. The institutional classification system, as mediated through the organization of a newspaper, generates what Cawelti (1970) has called a "formula" aesthetic, that is, "a highly conventionalized system for structuring cultural products."

All the constraints I have described, taken together, pressure newspaper photographers to take standard pictures for standard situations. Listen to how one photographer talks about how he works:

You have two choices. First, you can take the standard picture and forget about it. Or on the other hand, you can work very hard for the great picture. You crawl under man-holes and over buildings and you get the great shot. I'd say we've got a great deal of freedom in assessing what the news is and how to shoot the picture. But if they are not inclined to run my great shot, then I say, "Why bother?" I'll get the standard picture and forget about it.

Thus, while photographers have, in theory, a great deal of freedom to render their assignments in whichever way they choose, in practice they do not. I have repeatedly heard photographers say that they do not "go for the great picture" because the technology of news reproduction virtually cancels out varieties of imagery by placing limitations on the shapes, forms and black and white relationships in the photograph. Photographers also cite the column format of a newspaper as a limitation of the type of imagery they choose.

To summarize, we see that a variety of institutional and organizational constraints act in concert to steer photographers to take pictures that conform to an idealized version of a typical photograph for a typical situation. We have isolated the effects of mechanization, the division of labor and journalism qua institution and

can now see how these factors are determinants of stylistic conventions in news photography.

ADVERTISING

When an advertising agency produces an ad for a client that will appear in a print medium and, furthermore, the proposed ad calls for a photograph, the art director heading the campaign selects a photographer to do the job. The photographer he selects is usually a free-lance photographer who is in business for himself. Thus, the agency is the photographer's client, and if the photographer is successful, he has many clients.

However, owing to the organization of advertising, the photographer also inherits his client's client, the advertiser. This configuration imparts some interesting qualities to the relationship between the photographer, the agency and the advertiser, for it is the photographer's job to please two clients simultaneously. His economic dependency on the agency for future jobs constrains him to provide a variety of services, the least of which consists of his basic photographic skills. The photographer must supply visual expertise, must possess a cheerful and easygoing manner, must take orders in an accommodating way and must communicate special quality called "faith in the outcome" to people whose business is fraught with uncertainty. These latter attributes are essential in assuaging the anxieties of the advertisers and reducing the art director's tensions. To accomplish this, photographers often develop personal styles in which they tend to display great personal flair and an "arty" self to their dual clients.

"Doing a job" consists of providing the agency with a photographic version of the layout. A layout is a highly detailed, hand-drawn sketch of the proposed ad. The drawing indicates how the product is situated, positioned, and lit in its environment. Also indicated in the layout is the copy, the words and sometimes the type of lettering to be employed. The photographer's basic job is to translate the sketch into a photograph and give the pictures to the agency.

Providing a photographic rendition of the layout is not as easy as one might imagine. In the first place, a great deal of preparatory backstage work is required. If the ad is a fashion ad, models have to be selected. This is often done collectively by the art director and the photographer and the final selections often depend on the availability of the models. Locations have to be selected. Permission to shoot in many choice locations has to be obtained. Props and other background materials must be arranged and obtained, either through purchase or rental. Sometimes free-lance specialists are required to assist on a particular job and they must be hired by the photographer if they are not among his standing staff of darkroom assistants and secretaries. These might include a stylist, who advises and obtains objects to be placed in the setup, or a home economist who does with food, china, dinnerware and silverware what the stylist does with fabric, decorative ornaments and accessories. Other specialists might include a decorator, a cosmetologist, a hairdresser, a seamstress, mannequin dressers and graphic artists. If the shot calls for special technical equipment and photographic materials, they must be acquired. Preparation for the shooting is the first step in translating a sketch into a photograph. In fact, one can already see that the characteristics of the final photograph depend on the coordination of these collective activities (Becker, 1974).

The layout itself may present difficulties in translation and this sometimes produces conflict between the photographer and the art director.⁸ This, too, may be regarded as backstage work, since the advertiser has no knowledge of the negotiation between the photographer and art director. The following quotation provides an example of the type of technical difficulties that may arise in the translation of a drawing into a photograph. In this case, the issue is perspective:

The idea behind the ad was to have a taxi driver lie on a mattress. The way the art director drew the mattress was impossible to shoot, since the perspective was all off. The art director insisted that the photograph follow the layout to the inch. So I took the picture and then traced all the angles and the whole perspective was off. Then he finally gave in. Most of the time I work very close to the layout, but that dodo wanted it to the inch. Ordinarily, there is a little bit more flexibility.

The shooting itself is characteristically problematic for photographers in that a variety of difficulties, both anticipated and unanticipated, emerge at that time. Some problems, such as technical problems which could not have been anticipated prior to the actual shooting, are regarded by photographers as normal trouble (Cavan, 1972). I observed one occasion during which the shooting had to be stopped because the models, dressed in spring attire, were shivering in the near zero temperature at the beach location.

More problematic for a photographer are the difficulties that arise from the fact that he has dual clients who are often present at the shooting. The fact that they are present at the shooting in the first place is largely attributable to the photographer's need to please both clients. Anticipating this, he creates in his studio a "play" environment, supplied with rock music, drinks, pool or tennis tables, dart boards and sometimes beautiful people. Sometimes the noise level made by executives at play distracts the photographer and thus constitutes a form of interference. Photographer's secretaries and/or representatives are often present to take care of the client and to free the photographer to do his work.

The advertiser's presence at the shooting introduces difficulties for the photographer in ways other than described above. Since advertisers feel that photographers do not care about the product, they are present at the shooting to make sure that their product is photographed well. They tend to watch the photographer, police his action, ask a multitude of questions and seem to challenge his judgments repeatedly. They inquire as to whether the tiny creases in a garment will

⁸ The relationship between the art director and the photographer sometimes has conflicts of the same type described by Faulkner (1976) regarding Hollywood film composers and their clients. These conflicts are also found between the art director and his client, the advertiser.

be greatly visible in the final photograph or whether a slight shadow across the brand name will obliterate entirely the brand name in the final picture. In short, they worry that the photographer does not love their product as much as they do and they keep reminding him of that. At one shooting I attended, the advertiser dropped into the studio with the intention of seeing how the shooting was progressing and, as he phrased it, "to fool around a little." When he arrived and judged that the models were the wrong type and that the clothing was poorly displayed, he threw a tantrum and virtually stopped the shooting by placing himself directly in front of the camera. He yelled and screamed and eventually had to be cooled down by the agency's account executives who assured him he would not be charged for that day's expenses. Such instances, while perhaps not as dramatic as the foregoing account, typically occur in advertising photography and indicate that the demanding client does indeed wield a great deal of power.

Much of the character of the relationship between the photographer and his dual clients derives from the fact that this is essentially a bureaucratically organized professional client relationship which conjoins features of bureaucratic roles and professional roles in some peculiar and ambiguous way. Theoretically, the photographer should be treated as a professional who possesses esoteric knowledge and who can assist the client with his problem (Hughes, 1958:141). The fact that he is an outside expert, brought in especially for a specific job, should add to his professional status. Consequently we would expect that he would be accorded deference. This picture, however, is true only for top advertising and fashion photographers. For most advertising photographers, the situation is entirely different.

For most advertising photographers, the social organization of advertising has the net effect of chiseling away at the broad range of knowledge and expertise that the photographer brings with him. The photographer's contribution is virtually reduced to technical labor. The photographer is often given direct orders

by the art director and is told to photograph the models or objects the art director's way and not his way. When this occurs, the photographer's authority and expertise seem to evaporate entirely and, as his autonomy decreases, there is a concomitant reduction in the dimensions within which he is allowed to be creative on the job. In order to better understand how the photographer's role is narrowed to that of technician, we must explore client demands a little more to see how the institution of advertising impacts and shapes the social organization of photographer's work. Again, by looking at both social organization and institutional understandings, the reciprocal relationship between the two can be clarified.

Advertisers come to an advertising agency with preconceptions about what a good ad is. Their preconceptions are based on their previous experience with advertising agencies. Since they reason that all people who participate in the culture are exposed constantly to advertising and make evaluations about ads, many regard themselves as "expert amateurs" on advertising. They also come to the advertising agency with conceptions about what consumer's conceptual and perceptual orientations are and these, in part, are based on market research data. There is another factor, too, that shapes advertisers' conceptions of ads. The institution of advertising is similar to journalism in that there are basic stories. Many themes found in advertising—what may be called advertising stories—are institutionalized. A family breakfasting together or the "boys" having a beer after a game are examples of advertising stories. Even a close-up shot of a bottle of beer is a type of standard advertising picture. In short, advertising photography relies heavily on typical situations, typical themes and typical arrangements of people and/or products.

The advertiser expects the agency to provide a picture that partakes of some standard scenario, which is retrieved from the institutional stockpile of advertising stories. But the advertiser also expects the agency to add an original twist so that the final photograph will immediately capture the audience's attention and differentiate

the advertiser's product from the many competing products available on the market. In short, the advertiser wants an original standard picture, but the photograph must be conditionally original in the sense that the visual rendition does not undermine its impact as an advertising picture.

To comply with the advertiser's expectations, the ad agency's art director usually relies on the visual tradition of representationalism or pictorialism for the layout. Art directors seem to believe that advertisers will most readily approve of a straightforward pictorial rendition of their product. While art directors may add some visual elements that are striking, they generally leave it up to the photographer to provide the visual originality the advertiser wants. Thus, by the time the photographer is called for the job, the advertiser and the art director have settled on the type of visual imagery in the proposed ad and they expect the photographer to take their conception and breathe life into it, so to speak. Note that the photographer has not been asked to make suggestions or to comment on the selection of the imagery and its rendition. Instead, the photographer is called upon to supply his skills to turn a graphic design into a photograph. When the photographer makes an unsolicited suggestion, the photographer's comments are usually overruled by the coalition formed by the advertiser and art director. Only occasionally, photographers report, does the art director align himself with the photographer against the advertiser.

Given these constraints to begin with, the photographer's contribution is largely technical. Let us take an example of a "pretty girl displaying a product" picture. The photographer may vary the mood by using soft lighting and a soft focus lens and the resulting photograph will have a warm and intimate ambience. Or perhaps the photographer will use high contrast lighting and the result will be a cold and dramatic picture. Typically, photographers work within this functionally specific, narrow technical role that this form of social organization assigns to them. But it is within this narrow scope of the role of technical translator that advertising photographers define their creativity and

try to maximize their control. They accomplish this by doing as much technical variation as possible. I have seen photographers manipulate the image through a variety of technical means, including focus, lighting, distance, scale and distortion. At one shooting I observed, a photographer varied the camera angle by first 10 degrees and then 25 degrees and then changed the lighting.

It is not surprising then that a key stylistic feature of advertising photography is its visual unnaturalness as I described earlier. From the above discussion, we can see how the social organization of advertising itself is an important determinant of the type of visual imagery that we know as advertising photography, which consists essentially of a basic theme with much technical variation. The division of labor in advertising separates the conception of the image from its execution (see Read, 1965; Braverman, 1974; Rosenblum, 1978). This division of labor compromises the photographer's role as visual expert and reduces his work to technical labor.

On the institutional level, advertising photography is more heterogeneous than news photography. However, the directions in which it expands or the types of photographs it can absorb are essentially technically innovative and not thematically innovative. A review of the history of advertising photography will verify this point; a longitudinal content analysis will confirm this beyond a shadow of a doubt.

FINE ARTS PHOTOGRAPHERS

Fine arts photographers are not, strictly speaking, an occupational category and so, unlike news and advertising photographers, they are not constrained by the organizational features of production I have described in the other settings. As a consequence, the fine arts photographer retains his control over the technical and aesthetic aspects of the production of photographs. However, because there are no apparent structures that impinge upon the photographer's control over the work process, it does not mean that the fine arts photographer's choice of imagery is totally free from constraint. The major constraints that operate in the world of fine

arts photography are located at the distribution end. While they are more difficult to discern than those already identified, they operate forcefully nonetheless.

The fine arts photographer's freedom to select and produce imagery is narrowed when he takes into account the preferences and purchase pattern of gatekeepers in the market. To the extent that the photographer treats them as a reference group (Merton, 1957: 225-386; Shibutani, 1955; Turner, 1956) and orients his own work to their preferences, the photographer's work is constrained by the market. The marketplace, then, is a major source of control which feeds back to the photographer and exerts pressures on him to produce pictures in certain given directions and traditions, thus affecting his imagery.

For fine arts photographers there is a tension between their autonomy and the demands of the marketplace. This tension becomes dialectical when the pervasive institutional norm of originality in the arts conflicts with the contradictory demands of the art market. For most fine arts photographers, the selection of imagery is limited by social forces originating in the structure of the distribution channels. The imagery a photographer produces is often a result of the dialectical tension between his or her own personal imagery and institutionally rewarded imagery. In order to see these relationships clearly, I will first describe the fine arts photographer's high degree of control over the picture-making process and then describe the structure of the marketplace. This section on fine arts photographers will conclude with a discussion of the effects of the education-gallery-museum network on the photographer's selection of imagery.

The fine arts photographer determines for himself practically all aspects of the labor process, including the rhythm and pace of his work and the quality and quantity of his output. He determines what techniques will be employed, what equipment will be used. He controls each of the hundreds of little decisions that go into making the picture. Since there is generally no division of labor, he can make one decision with reference to how

he will do another (Dewey, 1934:45). In other words, the fine arts photographer can think in terms of an aesthetic totality and can apprehend various relations between all the steps, each of which are within his control. Thus, by not selling his services to a client and by not hooking into a structure where he is required to make pictures for his livelihood, the fine arts photographer retains his autonomy and control over the picture-making process.

To illustrate how much control over the work process fine arts photographers actually do have, let me furnish some examples of how they work. Fine arts photographers' work methods are as individual as each photographer. Unlike news and advertising, where many aspects of work are structurally and organizationally predetermined, fine arts photographers vary tremendously in their work methods. One photographer shoots every morning, develops his negatives every afternoon and prints every evening. Another photographer shoots all summer long and spends the long New England winters developing and printing his negatives. One photographer I interviewed returned to making pictures after a two-year "personal sabbatical" from photography. Another told me he had just chosen to enter a period of unproductivity in order to "replenish his juices." Personal work styles differ enormously, then, but the thread that links all the fine arts photographers in my sample is the virtual absence of the division of labor and the absence of extensive technology. With respect to the division of labor, every photographer did every step in the picture-making process by himself. He did everything from the routine tasks of mixing chemicals and taking temperature readings of solutions to the more demanding and complex work of printing the negatives. On occasion I found a division of labor; but, where it exists, its character is elementary. For instance, sometimes a famous photographer will hire his or her students to make multiple prints from one negative.

With respect to the technology of the photographic process, we saw that photography in news is highly mechanized. In

advertising, client demands steer photographers to produce technical innovations. In contrast, fine arts photographers dominate the technique. A remark made by one photographer applies to all the photographers I interviewed. He said, "Keep it simple," and in fact, every photographer had a simple setup. Of course darkrooms ranged in size and comfort and some photographers had more expensive equipment than others. But all in all, their darkrooms contained little gadgetry, no mechanized equipment and no fancy equipment which elaborated the basic photographic process. Perhaps one reason for this is the generally limited personal resources of the photographer. Nonetheless, fine arts photographers retain technical control over their work and, in this sense, continue to produce pictures in the craft tradition.

With respect to the selection of subject matter and its rendition, we saw how advertising and news photographers receive their assignments from other people. In other words, they are told what to shoot. In contrast, fine arts photographers' assignments are self-generated. They select what they want to shoot. In my sample, they reported various ways in which they became interested in some specific subject matter or theme. For some, new subject matter organically grows out of previous work. For others, preoccupation with an abstract idea, such as the relationship between lovers, propels the photographer to shoot specific subject matter. For another, it is a visual idea or a whim that engages the photographer and places him or her on a virtually uncharted course into a subject area.

Because the norms of originality and creativity are institutionally prevalent and are almost sacred in the art world, photographers are institutionally pressured to seek their own imagery and to render their subject matter in personally idiosyncratic ways. The following quotation expresses an idealized version of how free fine arts photographers feel with respect to their aesthetic choices:

I feel free. I can shoot cracks in the pavement or corkscrews or anything. I can look at it seriously and the result is that other

people pay attention to how I look at it. I shoot whatever I want, any time I want it.

One result of institutional encouragement of originality is the tremendous variation among the types of imagery that fine arts photographers produce. Photographers use all sorts of technical and aesthetic means to establish their own personal visions.

However, despite the wide variation of imagery produced by fine arts photographers, only a small percentage of it ever reaches the public, and it is here that the funneling role of the marketplace can be most clearly seen. The market is a powerful source of control that influences, in part, what photographers shoot and determines what the public eventually sees and defines as fine art photography.

The market is a highly competitive one, where there are many more photographers than there are outlets for exposure. Furthermore, the actual demand for purchase, while it has increased within recent years, does not approach the volume of the supply. Institutional gatekeepers, such as museum curators, key gallery directors, critics and important collectors, define the current fashion through their own exhibitions and purchases. These definitions change from time to time. When I began this research, photographs of poor people in both rural and urban settings were in vogue. When I finished, portraits of middle income, suburban families, often depicted as participating in family ritual, were popular. Another type of subject matter that was popular when I first began was the photographer's own family rendered in classically beautiful ways. Now the "snapshot aesthetic" is fashionable. In short, gatekeepers can make a trend.

Economic constraints imposed by the competitive market structure of fine arts photography exert pressure on gallery directors and museum curators to seek and find a winner. They will mobilize their resources and stake their reputations on a winner. If other people in the art world, especially critics, also agree that the photographer is a winner, the backers will eventually reap the financial and prestige

benefits of their initial wisdom. But, in actuality, museums and galleries rarely go to bat for a new photographer or show his work, for there is little financial success in that. Rather, upcoming photographers are shown in group shows. More often than not, the gallery will show the new work in some of its exhibition space while elsewhere in the gallery the old reliable masters are on permanent exhibition. Buyers might enjoy new work but they buy established photographers whose work has already been catalogued and classified under some school or tradition. Thus, the economics of the art world has the latent consequence of maintaining the dominance of several photographic traditions, i.e., the West Coast landscape school. The influence of tradition or school is powerful in the art world. It can be said that they constitute a loose category system of the sort described earlier for news and advertising. Of course, the categories are much more open in the fine arts and there is no enforcement structure or threat of loss forcing photographers to comply with any one tradition. But the market rewards photographers who do work within one of several conventional traditions. The market, then, constrains photographers to select imagery within certain given directions. Here is an example of how these structural factors are seen by one especially articulate photographer:

I've won a Guggenheim, which is the highest award for a fine art photographer. My work hangs in museums. But don't think that the museums are so noble and pure and sacred. X curator, he influences photographers' work, for good or for ill. The photographer develops a client relationship with the museums. X has an aesthetic that can be described: he likes photographs taken with large cameras, that are posed and that have a 19th century character to them. I personally feel that it's dangerous when an institution begins shaping the kind of pictures photographers make. And then Y is homosexual and he doesn't like pictures of women. And then Z has thoroughly academic and formalistic tastes. The museums have encouraged a preoccupation with the neurotically personal and the aridly formal at a time when we are surrounded by an atmosphere of tremendous social agitation and polariza-

tion. My work is involved with social meanings and there is no museum or gallery that wants to show or buy my work now.

Photographers who do not produce imagery which conforms to prevailing definitions of photography are faced with the problem of how to get their work shown. For instance, a number of San Francisco and San Diego photographers see the art world as an institution of social and aesthetic control. Ideologically they resist the hegemony of powerful gatekeepers and the prevailing definitions of photography as Art by rejecting the definition of the single image as an autonomous art object. In their view, a photograph has no meaning in isolation. Rather, its meaning is constructed through juxtaposition with other imagery. Since many of them have been influenced by semiology, they juxtapose the image with text but treat the text as part of the image, rather than treat it as ancillary to the image. Their definition of photography finds no sympathy or support among gatekeepers. Consequently, these photographers have to create their own alternative systems for the distribution of their work. This case shows the extent to which prevailing definitions filter down to regional and local galleries and the extent to which the economics of the art market upholds institutional definitions of pictorial photography, a tradition borrowed from painting. In sum, the market may be seen as a source of control through the creation of boundaries, definitions and meanings. Constraints on photographers to produce work within a few established traditions arise from market allocation of rewards. In other words, the market determines definitions of fine arts photography by virtue of what is selected for inclusion.

Another source of control over the types of imagery photographers may choose is the effect of the education-gallery-museum network system in the world of fine arts photography. There are linkages between gatekeepers and important university and college photography departments. Students are structurally channeled into working within the tradition of their mentors and are often rewarded for that by the connections their

mentor makes with his or her agent or gallery. Thus, the system of sponsorship can be seen as an extension of the marketplace in that it provides access and is a system of privileged recruitment. While this social arrangement furnishes a mobility path for upcoming young photographers, at the same time it tends to perpetuate the dominance and continuity of several photographic traditions. For instance, one gallery director reported to me that she could identify whom a student studied with simply by looking at the student's portfolio.

The task of explaining why there is wide variation of imagery in fine arts photography is not an easy one. On the one hand, it could be argued that the multiplicity of imagery in fine arts photography is the result of normative pressure to be original. We would then expect that fine arts photography would consist of multitudes of individualistic images, distinguished by personal symbolic systems, introverted complexities and idiosyncratic renditions. In other words, we would expect fine arts photography to be a residual category that scoops up all the personal and unclassifiable pictures that do not fit under other categories of photography. We could also explain the wide variation in imagery by saying that, since fine arts photographers retain their autonomy over the picture-making process, they can produce entirely individualistic images.

But the fact is that fine arts photography is not a residual category and it does not have unlimited capacity to absorb all types of imagery. While its boundaries may be more elastic than those of news and advertising photography, the boundaries are there nonetheless. Also, the fact that fine arts photography is internally differentiated with respect to schools or traditions is another dimension that restricts the types of imagery admitted into fine arts photography. This suggests that the economics of the art market modify, temper and sometimes contradict the idealized version of the unrestrained and original artist. The strength of the market's influence is evident when we realize how few photographic traditions are defined as fine art. The market's influence is also evident when we consider how mar-

ket demands steer photographers to produce imagery in certain directions—imagery which often partakes of conventional traditions. Of course, there are occasional exceptions. Some careers, in fact, are begun by an outrageous violation of convention and tradition. But these cases are rare indeed. Generally, photographers get shows when their work has been defined as an extension of some already existing tradition in fine arts photography. In sum, social structural features influence the production of imagery in the world of fine arts photography; but, in contrast to news and advertising, the structural sources of convention are generated primarily through the distribution systems, rather than through the organization of production.

SUMMARY AND CONCLUSION

The preceding review of styles in photography and the social organization of the settings in which these pictures are made shows that there is an association between type of photographic style and type of social organization. This association is the basis for my claim that style is not exclusively a product of an autonomous set of shared understandings, but is at least partly determined by social organization. We saw that conventions or shared agreements can arise from, and in response to, such structural and organizational features as the division of labor, technology, the structure of the market and client relationships. Each of these factors is a source of control and each exerts different constraints on a photographer's work, the photographic process and the final photograph. Thus, photographers have different types of control over different dimensions of the picture-making process in each setting. Systematic observation of such differences helps us understand more precisely the mechanisms by which the organization of work "shapes the nature of work itself" (Friedson, 1975:94) and, ultimately, the relationships between the worker's autonomy and discretion (Thompson, 1967:117-31) and stylistic outcomes.

This study suggests some general propositions for further research. One impli-

cation is that the greater the rationalization of the work process, the more homogeneous the style and the less its capacity to absorb variation. One, of course, may argue that rationalization of the work process is associated with the monopolistic concentration of capital production and that, together, both are responsible for the development of a narrow vocabulary of stylistic conventions. The examination of styles of automobiles would be an excellent research site to determine the influence of these processes on style.

One way to approach the examination of the relationship between rationalization of the work process and style is to focus on the impact of technology. Technology, taken in the most basic ecological materialist sense, affects style and stylistic change. As Ivins (1953) has shown, the development of the engraving technology, which permitted pictorial images to be repeated exactly, effect major changes in the conventions, symbolism and syntax of the graphic arts. Jazz is another example. Hennessey (1975) points out that jazz was "nationalized," that is, local and idiosyncratic musical conventions were either dropped or incorporated when radio broadcasting and recordings became widely available. This same argument, I suspect, can be applied to the case of newspaper photography when the wire services hooked up virtually all the newspapers in the United States. We might study technological innovations, ranging from the invention of new musical instruments to television, to investigate their impact on stylistic forms.

The client's influence on the final form of musical film scores has been described by Faulkner (1976). In addition, there is ample evidence in the history of art that the artist's patron exerted tremendous influence in determining how paintings turned out. We might explore other contemporary social situations, especially those which are bureaucratically organized, in which the patron role is not as clearly defined as in past times but, nonetheless, exerts similar influences. The influence of granting agencies on publicly supported educational television might be an interesting place to start.

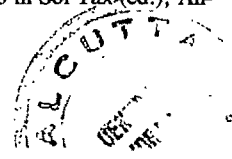
The numbers and types of styles that develop under different market conditions is important to explore systematically. Despite its difficulties, the Peterson and Berger (1975) study of musical styles represents one way to sociologically analyze styles of cultural products. Similarly Martorella's (1977) study of the determinants of opera repertoire also shows the connection between economics and characteristics of cultural products, as well as their social distribution.

Finally, another implication of this study is that complex and diverse imagery develops when remote institutional influences, rather than proximate organizational ones, dominate. To examine this, we need only turn to the history of art to compare styles under conditions of government support, sponsorship and distribution with styles which developed in nonregulated, competitive markets. One important market variation to study concerns the number of styles that can develop and flourish despite the hegemony of an official gatekeeping class. For instance, a cross-national study which looks at homogeneity and diversity among art styles as a function of degree of centralization of the gatekeeping class would be an important extension of the work already done by the Whites (1965).

The sociological study of style is an important avenue of inquiry which can contribute a great deal to our understanding of material culture and its relationship to social processes. Scholars in many other disciplines are now using stylistic analysis as an adjunct to their usual methods. Art historians (Gombrich, 1960; Schapiro, 1962), anthropologists (Kroeber, 1948), linguists (Fowler, 1975), structuralists in a variety of disciplines (de George, 1972; Ehrmann, 1970) and others attempt to answer such questions as "Why does this piece look (or sound) the way it does?" and "Why is this object formed the way it is?" Sociologists can provide a *social* answer to such questions. I have attempted to sketch out some possibilities for a sociological approach to the analysis of styles. Clearly, a sociological approach to this underinvestigated area is an important enterprise.

REFERENCES

- Baxandall, Lee
1972 *Radical Perspectives in the Arts*. Middlesex: Pelican Books.
- Becker, Howard S.
1974 "Art as collective action." *American Sociological Review* 39:767-76.
- Bennett, H. Stith
1972 *Other People's Music*. Ph.D. dissertation, Sociology Department, Northwestern University, Evanston.
- Berger, Peter and Thomas Luckmann
1966 *The Social Construction of Reality*. New York: Doubleday.
- Bernstein, Basil
1971 *Class, Codes and Control*. London: Routledge.
- Braverman, Harry
1974 *Labor and Monopoly Capital*. New York: Monthly Review Press.
- Burns, Elizabeth
1972 *Theatricality: A Study of Convention in the Theatre and in Social Life*. New York: Harper Torchbooks.
- Burns, Elizabeth and Tom Burns
1973 *Sociology of Literature and Drama*. Middlesex: Penguin.
- Cavan, Sherri
1972 *Hippies of the Haight*. St. Louis: New Critics Press.
- Cawelti, John
1970 *The Six-Gun Mystique*. Bowling Green: Bowling Green University Popular Press.
- Croce, Benedetto
1968 *Aesthetic*. New York: Noonday Press.
- Crozier, Michael
1964 *The Bureaucratic Phenomenon*. Chicago: University of Chicago Press.
- de George, Richard and Fernande de George
1972 *The Structuralists from Marx to Levi-Strauss*. New York: Doubleday.
- Dewey, John
1934 *Art as Experience*. New York: Capricorn Books.
- Ehrmann, Jacques
1970 *Structuralism*. New York: Doubleday.
- Faulkner, Robert R.
1976 "Dilemmas in commercial work: Hollywood film composers and their clients." *Urban Life* 5:3-32.
- Fields, Karen
1976 *The socialization of the artist in the United States*. Ph.D. proposal, Anthropology Department, Stanford University, Stanford.
- Fischer, Ernest
1963 *The Necessity of Art; A Marxist Approach*. Baltimore: Penguin.
- Fowler, Roger
1975 *Style and Structure in Literature: Essays in the New Stylistics*. Ithaca: Cornell University Press.
- Frederico, Ronald C.
1974 "Recruitment, training and performance: the case of ballet." Pp. 249-61 in Phyllis L. Stewart and Muriel G. Cantor (eds.), *Varieties of Work Experience: The Social Control of Occupational Groups and Roles*. New York: Wiley.
- Friedson, Eliot
1975 *Doctoring Together*. New York: Elsevier.
- Gombrich, Walter
1960 *Art as Illusion*. New York: Pantheon Books.
- Griff, Mason
1964 "The recruitment of the artist." Pp. 64-91 in Robert N. Wilson (ed.), *The Arts in Society*. Englewood Cliffs: Prentice-Hall.
- Hecht, Chandra Mukerji
1971 *Flashes and Pans: the Education of Student Film Makers*. Ph.D. dissertation, Sociology Department, Northwestern University, Evanston.
- Hennessy, Thomas J.
1975 "The nationalization of jazz, 1929-1935: pattern for popular culture." Unpublished paper.
- Hughes, Everett C.
1958 *Men and Their Work*. New York: Free Press.
- Ivins, Williams M., Jr.
1953 *Prints and Visual Communication*. Cambridge, Ma.: MIT Press.
- Kealy, Edward
1974 *The Real Rock Revolution: Sound Mixers, Social Inequality and the Aesthetics of Popular Music Production*. Ph. D. dissertation, Sociology Department, Northwestern University, Evanston.
- Kroeber, Alfred L.
1948 *Anthropology: Culture Patterns and Processes*. New York: Harcourt, Brace.
- Langer, Suzanne
1942 *Philosophy in a New Key*. New York: Mentor Books.
- Martorella, Roseanne
1977 *The Relationship between Box-Office and Repertoire: The Case of Opera*. *Sociological Quarterly* 18:354-66.
- Merton, Robert K.
1957 *Social Theory and Social Structure*. New York: Free Press.
- Meyer, Marshall W.
1971 *Structures, Symbols, and Systems*. Boston: Little, Brown.
- Molotch, Harvey and Marilyn Lester
1974 "News as purposive behavior: on the strategic use of routine events, accidents and scandals." *American Sociological Review* 39:101-12.
- Peterson, Richard A. and David G. Berger
1975 "Cycles in symbol production: the case of popular music." *American Sociological Review* 40:158-73.
- Read, Sir Herbert
1965 *Icon and Idea*. New York: Schocken.
- Rosenblum, Barbara
1978 *Photographers At Work: An Empirical Study in the Sociology of Aesthetics*. New York: Holmes and Meier. In press.
- Schapiro, Meyer
1962 "Style." Pp. 278-303 in Sol Tax (ed.), *Art*.



- thropology Today. Chicago: University of Chicago Press.
- Schutz, Alfred
 1964 Collected Papers. 2 Vols. The Hague: Martinus Nijhoff.
- Schuyler, Philip
 1976 "Master and pupil: music education in Morocco." Unpublished paper. Department of Ethnomusicology, University of Washington, Seattle.
- Shibutani, Tamotsu
 1955 "Reference groups as perspectives." *American Journal of Sociology* 60:562-9.
- Smades, Lisa
 1976 "Constructing dance technique: the use of praise and correction to define right and wrong." Unpublished paper. Sociology Department, Stanford University, Stanford.
- Spurrell, Margaret
 1974 The Formation of a Student Aesthetic within a University Theatre Department. Ph.D. dissertation, Department of Sociology, Northwestern University, Evanston.
- Thompson, James
 1967 *Organizations in Action*. New York: McGraw-Hill.
- Tuchman, Gaye
 1973 "Making news by doing work: routinizing the unexpected." *American Journal of Sociology* 79:110-31.
- Turner, Ralph H.
 1956 "Role taking, role-standpoint and reference group behavior." *American Journal of Sociology* 61:316-28.
- Weber, Max
 1968 *Economy and Society*. New York: Bedminster Press.
- White, Cynthia and Harrison C. White
 1965 *Canvasses and Careers*. New York: Wiley.
- Wolff, Janet
 1975 *Hermeneutic Philosophy and the Sociology of Art*. London: Routledge.

COMMENT

FACETS AND FLAWS OF HOPE'S DIAMOND MODEL*

(COMMENT ON HOPE, ASR JUNE, 1975)

Keith Hope's (1975) diamond model of status inconsistency and social mobility effects appears to offer a method for reestablishing the empirical utility of such effects in the face of a decade of largely negative evidence (see Jackson and Curtis, 1972).¹ This comment reviews and clarifies Hope's argument and its relation to prior theory and research on status inconsistency. This analysis suggests that Hope in no way challenges the extant body of negative evidence, but rather proposes a new definition of inconsistency which neither accords with the definition shared by sociologists at least since Lenski's (1954) seminal work nor has merit in its own right. Further Hope's new definition assumes the existence of a unitary dimension of general social status—an assumption which is largely unfounded and certainly inconsistent with recent trends in the study of social stratification. Thus, though some of the concerns that led Hope to propose his diamond model are valid, the model is not useful for the study of status inconsistency, nor probably for the study of stratification more generally.

Points of Agreement

Hope proposes his model as an alternative to the square additive and linear additive models for analyzing status inconsistency effects (see below) developed by Lenski (1964), Duncan (1966) and Blalock (1966; 1967). Hope rightly views most applications of these approaches as atheoretical for failing to specify the nature of the additive (i.e., status) and interactive (i.e., inconsistency) effects involved. Such theoretic-

cal ambiguities bias the resulting empirical analyses toward models positing only additive status effects and against inconsistency effects (Hope, 1975:326,330,332,339). He also is correct that, beyond these problems of theoretical specification, the multiple regression procedures suggested by Duncan and Blalock constitute a very conservative test for inconsistency effects, though this conservatism was intended (see Duncan, 1966; Hope, 1975:329,336). Thus Hope is rightly discontented with the existing status inconsistency literature, but his proposed remedy is more dangerous than the ills it seeks to cure.

Prior Approaches to Inconsistency Effects

The flaws in Hope's diamond model are best seen in light of the prior history of status inconsistency effects. Table 1 presents both a general paradigm and a hypothetical example which illustrate the basic theoretical and methodological problems of inconsistency research. Such research seeks to assess the separate and joint effects on a dependent variable (Y) of two status dimensions (A and B) on which individuals are ordered (e.g., occupational prestige and education). Cells *a* through *i* of the table would contain mean scores (or percentages) on the dependent variable for persons with particular combinations of status levels (e.g., cell *a* contains people low on both A and B). Persons in cells *a*, *e*, and *i* are status consistent, while those in cells *b*, *c*, and *f* are characterized by one pattern of inconsistency ($B > A$) and those in cells *d*, *g*, and *h* by the opposite type of inconsistency ($A > B$). At the bottom of Table 1 hypothetical mean scores are entered in the general table. If one assumes equal numbers of persons in each cell, the average scores of consistents (mean = 0), $A > B$ inconsistent (mean = -4), and $B > A$ inconsistent (mean = +4) differ considerably from each other; that is, there is apparent evidence of inconsistency effects.²

* I am indebted to Richard Campbell, Richard Cohn, Charles Hirschman, Wendy Fisher House, and Alan Kerckhoff for comments on earlier drafts, and especially to Mark Evers and an anonymous reviewer for their insights into statistical aspects of the paper.

¹ As Hope (1975) and others (e.g., Jackson and Curtis, 1972) suggest, analyses of inconsistency and mobility effects involve the same theoretical and methodological issues. For clarity and brevity, this comment discusses only status inconsistency effects, but the arguments generalize directly to mobility effects.

² If all inconsistent in Table 1 regardless of type are compared with the consistents, the inconsistency effects disappear. If the cell *n*'s were not equal, however, as they usually are not in actual research, the mean score of all inconsistent would differ from that for consistents as well as the two types of inconsistent differing from each other and from the consistents.

Table 1. Conventional Model for Analysis of Status Inconsistency Effects

		Status B			
		(Low)	(Med.)	(High)	Row
Status A		0	1	2	Mean
A. General Paradigm					
(Low) 0	<i>a</i>	<i>b</i>	<i>c</i>	$\bar{Y}_0.$	
(Med.) 1	<i>d</i>	<i>e</i>	<i>f</i>	$\bar{Y}_1.$	
(High) 2	<i>g</i>	<i>h</i>	<i>i</i>	$\bar{Y}_2.$	
Col. Mean	$\bar{Y}_{.0}$	$\bar{Y}_{.1}$	$\bar{Y}_{.2}$	$\bar{Y}_{..}$	
B. Hypothetical Example					
(Low) 0	+2	+4	+6	+4	
(Med.) 1	-2	0	+2	0	
(High) 2	-6	-4	-2	-4	
Col. Mean	-2	0	+2	0	

Yet these three groups also differ in their standing on status A and status B. On the average, A>B consistents are higher on A and lower on B than either consistents or B>A inconsistent. Inconsistency effects, however, are clearly meant to reflect more than the effects of the separate status dimensions, and must be assessed after controlling for persons' standings on status A and status B as suggested in Lenski's (1954:405-6, emphasis added) seminal hypothesis: "individuals characterized by a low degree of status crystallization differ significantly in their political attitudes and behavior from individuals characterized by a high degree of crystallization, *when status differences in the vertical dimensions are controlled.*"

The cell means in Table 1 were, in fact, generated by adding the effects of status A and status B (i.e., adding the row and column means applicable to each cell). Thus, these data contain no inconsistency effects in the sense that Lenski and others conceived of such effects. Yet comparing mean scores of consistent and inconsistent groups, which was the basic procedure in studies prior to 1964, yields apparent inconsistency effects virtually anytime the dependent variable is affected by either or both of the status dimensions, regardless of whether there is any effect of inconsistency per se.

In the mid-1960s this methodological flaw was revealed by Mitchell (1964), recognized and indeed clarified by Lenski himself (1964), and fully elaborated by Duncan (1966) and Blalock (1966; 1967). The prior work of Lenski and others had not adequately controlled "status differences in the vertical dimensions," and hence had identified as inconsis-

tency effects what were nothing more than additive effects of the separate status dimensions. Responding to Mitchell, Lenski (1964:326) clarified ambiguities in his earlier theory and methods:

in statistical terms, my hypothesis is that inconsistencies in status generate an interaction effect on variables that are symptomatic of stress . . . and that the simple additive model that students of stratification have traditionally used is incapable of describing relationships between two status variables and a variable measuring stress.

Duncan (1966) and Blalock (1966; 1967) presented multiple regression equations which estimate and test inconsistency effects while adequately controlling for effects of the separate stratification dimensions:

$$\hat{Y} = b_1A + b_2B; \quad (1)$$

$$\hat{Y} = b_1A + b_2B + b_3AB. \quad (2)$$

Equation (1) estimates the additive effects of the two status dimensions (A and B) on the dependent variable (Y), while equation (2) estimates these effects plus the effect of some interactive (nonadditive) combination of A and B. The size and significance of the inconsistency effects is estimated by the additional variance (R^2) explained by equation (2) compared to equation (1). These procedures would reveal *no* inconsistency effect, for example, in Table 1.

Hope's Flawed Diamond Model

Hope observes that rotating an inconsistency table such as Table 1 yields the diamond-shaped Table 2, hence the term

Table 2. Diamond Model for Analysis of Status Inconsistency Effects

Status (A+B)	Inconsistency (A-B)					Row Mean
	2	1	0	-1	-2	
A. General Paradigm						
(Low) 0			<i>a</i>			$\bar{Y}_{0.}$
1		<i>d</i>		<i>b</i>		$\bar{Y}_{1.}$
2	<i>g</i>		<i>e</i>		<i>c</i>	$\bar{Y}_{2.}$
3		<i>h</i>		<i>f</i>		$\bar{Y}_{3.}$
(High) 4			<i>i</i>			$\bar{Y}_{4.}$
Col. Mean	$\bar{Y}_{.2}$	$\bar{Y}_{.1}$	$\bar{Y}_{.0}$	$\bar{Y}_{.-1}$	$\bar{Y}_{.-2}$	$\bar{Y}_{..}$
B. Hypothetical Example						
(Low) 0			+2			2
1		-2		-4		1
2	-6		0		+6	0
3		-4		+2		-1
(High) 4			-2			-2
Col. Mean	-6	-3	0	+3	+6	

"diamond" model. The cells in Table 2 are the same as in Table 1, but the vertical and horizontal dimensions have changed—the vertical dimension (rows) of the diamond represents the sum of the separate status variables ($A + B$), while the horizontal dimension represents the difference ($A - B$) between these two variables, i.e., the degree of status inconsistency. The row means of the diamond, then, represent the effect of a person's total or general status level on the dependent variable, while the column means represent the effect of different degrees and directions of inconsistency or mobility. The lower portion of Table 2 reveals an apparent inconsistency effect which is even stronger than that derived from the original procedure of comparing consistent and inconsistent, but is clearly contrary to the results (i.e., no inconsistency effect) which the later Lenski or Duncan-Blalock procedures yield for the same data.

Hope (1975:326-7) also translates his diamond model into a regression framework. Estimates and tests of inconsistency and mobility effects are to be derived from the following equations:

$$Y = b'_1 (A + B); \quad (3)$$

$$Y = b'_1 (A + B) + b'_2 (A - B). \quad (4)$$

The two terms in equation (4)—($A + B$) and ($A - B$)—are merely algebraic expressions for the status and inconsistency dimensions of the diamond table.³ Hope posits that an inconsistency effect is demonstrated if equation (4) explains significantly more variance in Y than equation (3). This procedure would yield a status inconsistency effect for the data in Tables 1 and 2, whereas the Duncan-Blalock regression procedures would not.

Hope appears to have pulled a methodological coup; it appears he has empirically documented inconsistency effects in data initially thought to contain such effects but subsequently shown to contain only additive effects of the two separate stratification dimensions. In fact, Hope's procedures reveal no new properties or information in the data. As

³ It is worth noting (as Hope does not) that adding and subtracting A and B requires that they be in the same metric. But stratification variables are usually not in the same metric; they can be converted to the same metric only through arbitrary assignment of values as in Table 2 or through standardization. The former procedure is clearly arbitrary while the latter makes analyses incomparable across populations or samples characterized by different distributions (or variances) of A or B . In short, besides its other flaws, Hope's model necessitates use of unreasonable and/or undesirable procedures to equate the metrics of the status variables.

Hope is aware, equation (4) is merely a transformation of equation (1) or the basic additive model. Equations (1) and (4) yield identical R^2 's and identical predicted values of Y . Further, the metric regression coefficients in equation (4) are derivable from those in equation (1); e.g., $b'_2 = b_1 - b_2/2$.⁴ That is, Hope's diamond model yields evidence of an inconsistency effect whenever the two status variables involved have unequal effects on a dependent variable, that is, whenever $b_1 \neq b_2$ in equation (1).⁵ For example, if income is strongly related and education only weakly related to amount spent on entertainment by a consumer, Hope's procedures yield an inconsistency effect on spending for entertainment. Obviously, Hope has a very different conception than previous analysts of what constitutes status inconsistency and an effect thereof.

In sum, Hope's quarrel with the approach of Duncan, Blalock and Lenski himself is not really methodological or empirical but rather theoretical and conceptual. Equations (1) and (4) are simply different ways to represent a single empirical reality: (a) as separate stratification variables which combine additively to predict a dependent variable (Y), or (b) as a summative dimension of general social status and a difference between the stratification variables composing it, with the general dimension and the difference combining additively to predict (Y). Thus, Hope departs conceptually from Lenski, Duncan, and Blalock in two related ways: (1) to control for "status differences in the vertical dimensions" is taken to mean controlling for a single composite dimension of general social status rather

⁴ The following algebra shows why the metric coefficients in equation (4) are derivable from those in equation (1):

$$\begin{aligned} \hat{Y} &= b'_1 (A + B) + b'_2 (A - B), \\ &= b'_1 A + b'_1 B + b'_2 A - b'_2 B, \\ &= (b'_1 + b'_2) A + (b'_1 - b'_2) B. \end{aligned}$$

But it must be true that $b_1 = (b'_1 + b'_2)$ and $b_2 = (b'_1 - b'_2)$.

Hence $b'_1 = b_1 - b'_2$ and $b'_2 = (b'_1 - b'_2) - b'_2$.

Hence $b'_2 = b_1 - 2b'_2$ or $2b'_2 = b_1 - b_2$ or $b'_2 = \frac{(b_1 - b_2)}{2}$.

Similarly we find $b'_1 = \frac{(b_1 + b_2)}{2}$.

I am indebted to Mark Evers for first showing me these basic derivations.

⁵ Hope (1975:327) allows the status variable ($A + B$) in equation (4) to be a weighted sum of A and B ($w_1A + w_2B$), with the weights unequal. As these weights w_1 and w_2 approach b_1 and b_2 of equation (1), respectively, Hope's inconsistency effect— b'_2 in equation (4)—diminishes, becoming zero if $w_1 = b_1$ and $w_2 = b_2$. Hence Hope precludes the use of equation (1) for arriving at the appropriate weights.

than for the set of separate additive effects of the vertical dimensions; and (2) a status inconsistency effect is defined as two stratification variables having unequal impacts on a dependent variable, rather than as a statistical interaction between the two variables.

Hope's position on both these points must be rejected on substantive and theoretical grounds. Hope's redefinition of status inconsistency negates all previous understanding of the concept and makes it so commonplace as to be trivial. Virtually all status variables relate somewhat differently to a given dependent variable. If we accept Hope's definition of status inconsistency, almost every dependent variable which is affected by any status variable would also be affected by status inconsistency. The above example of such a status inconsistency effect on consumer spending would be one of millions of such effects. It seems much more sensible to say, for example, that income has a stronger effect than education on consumer spending, than to relabel such phenomena as "status inconsistency effects."

Hope's inconsistency term ($A - B$) in equation (4) could not be estimated in the same regression equation in which A and B were entered as separate variables. Thus, Hope's theoretical definition of inconsistency and his methods of estimating its empirical effects necessitate that we control for status differences in the vertical dimensions by controlling a single composite dimension of general social status. Yet Hope presents only two arguments for this strategy, and neither is convincing. Hope argues (1975:326-7) that Lenski (1954) originally intended that analyses of status inconsistency control only for a single composite vertical dimension rather than for the combined effects of the separate status dimensions. This interpretation is, however, not consistent with Lenski's (1954:406, emphasis added) original statement about the need to control for "differences in the vertical dimensions" of status. It also denies Lenski the right to clarify his theory as he did in 1964 (see above) and implies that we make no progress in refining theories over time. Secondly, Hope (1975:342) argues that concepts and measures of a single composite dimension of social status are "analytical," while treatments of stratification in terms of a series of separate dimensions are merely "descriptive." But he does not explain what he means by analytical and descriptive. It appears he prefers the composite status dimension because it is derived (e.g., by statistical procedures such as factor analysis) rather than directly observed. Whether a variable is directly observable or is a composite of several other directly observed variables says nothing

about its scientific value, which lies, rather, in its ability to help us conceptualize, understand, and predict empirical events. In fact the striking advances in stratification research and theory over the past ten to fifteen years have stemmed from efforts to disaggregate social stratification into its component dimensions and to understand the causal interrelations among these dimensions and between them and other social phenomena (e.g., Blau and Duncan, 1967; Hauser, 1973; Lenski, 1966). Composite dimensions of general social status appear to have little scientific value in research on either status inconsistency or stratification more generally. The same is true of Hope's diamond model.

James S. House
Duke University

REFERENCES

- Blalock, Hubert M., Jr.
1966 "The identification problem and theory building: the case of status inconsistency." *American Sociological Review* 31:52-61.
1967 "Status inconsistency and interaction: some alternative models." *American Journal of Sociology* 73:305-15.
- Blau, Peter M., and Otis Dudley Duncan
1967 *The American Occupational Structure*. New York: Wiley.
- Duncan, Otis Dudley
1966 "Methodological issues in the analysis of social mobility." Pp. 51-97 in Neil J. Smelser and Seymour M. Lipset (eds.), *Social Structure and Mobility in Economic Development*. Chicago: Aldine.
- Hauser, Robert M.
1973 "Disaggregating a social-psychological model of educational attainment." Pp. 255-84 in Arthur S. Goldberger and Otis Dudley Duncan (eds.), *Structural Equation Models in the Social Sciences*. New York: Seminar Press.
- Hope, Keith
1975 "Models of status inconsistency and social mobility effects." *American Sociological Review* 40:322-43.
- Jackson, Elton F., and Richard F. Curtis
1972 "Effects of vertical social mobility and status inconsistency: a body of negative evidence." *American Sociological Review* 37:701-13.
- Lenski, Gerhard
1954 "Status crystallization: a non-vertical dimension of social status." *American Sociological Review* 19:405-13.
1964 Comment. *Public Opinion Quarterly* 28:326-30.
1966 *Power and Privilege: A Theory of Social Stratification*. New York: McGraw-Hill.
- Mitchell, Robert E.
1964 "Methodological notes on a theory of status crystallization." *Public Opinion Quarterly* 28:315-25.

ITEMS (Continued)

issue of the *International Journal of Health Services* entitled *Social Services and the State: Class Conflict and Accumulation* (forthcoming). EDWIN D. JOHNSON is a Ph.D. Candidate at the University of Wisconsin, Madison. His research centers on the impact of organized labor on state economic policies and the influence of local and regional power configurations on the distribution of federal grants and programs.

■ MICHAEL ARMER (Consequences of Psychological Modernity) is Associate Professor in the Department of Sociology at Indiana University. He is extending his research on the measurement and behavioral consequences of psychological modernity to the U.S. and Nigeria as well as Costa Rica. He also is studying the influence of educational and occupational experience on personality in developing societies. LARRY ISAAC is Assistant Professor of Sociology at Florida State University. His most recent work has focused on collective political violence and social change. He is involved in research on the relationship of fiscal behavior of advanced industrial states to structural features and economic prosperity.

■ JOHN D. BALDWIN (Verstehen and Erklären) is Associate Professor of Sociology at the University of California, Santa Barbara. JANICE I. BALDWIN is a Lecturer at the University of California, Santa Barbara. The Baldwins are conducting behavioral research on nonhuman primates and humans. Their research focuses special attention on exploration, play and creativity.

■ DIANA B. DUTTON (Low Use of Health Services by the Poor) is Assistant Professor of Sociology and Health Services Research in the School of Medicine, Stanford University. Her research interests include the quality of primary health care in alternative health delivery systems based on (clinically assessed) children's health disorders and the political sociology of the biomedical research decision-making process and its consequences.

■ R. JAY TURNER (Social Factors in Psychiatric Outcome) is Professor of Sociology and of Epidemiology and Preventive Medicine, and Director of the Health Care Research Unit at the University of Western Ontario, London, Ontario. His current research efforts include further work on the

influence of social factors on the occurrence, course and outcome of psychiatric disorder, collaborative studies (with Carl Grindstaff) in the area of teenage pregnancy, and a consideration of the association between social class position and illness susceptibility. JOHN W. GARTRELL is Associate Professor in the Department of Sociology, University of Alberta. Besides continuing writing on the influence of social factors in mental illness, he is currently involved in a study of alienation from work (with Peter Archibald) and is completing an analysis of development and inequality in Agrarian Indian Communities.

■ BLAIR WHEATON (Sociogenesis of Psychological Disorder) is Assistant Professor of Sociology at Yale University. His research interests include attributional explanations for social causation of psychological disorder, the cross-cultural measurement of life stress, and the derivation of nonlinear models of the relationship between socioeconomic background, varying patterns of educational and early occupational socialization, and achievements in later adult life.

■ THEODORE L. REED (Organizational Change in the American Foreign Service) is Assistant Professor of Sociology at Temple University. His publications include *A Theory of Public Bureaucracy: Politics and Personality in the Department of State* (Donald P. Warwick in collaboration with Reed and Marvin Meade, Harvard University Press, 1975). He is coauthor (with David Anderson) of a forthcoming book entitled *Race in America: A Historical and Institutional Perspective*. He is continuing research on the political economy of complex organizations. Currently he is doing research for a monograph on the historical development of the bureaucratic state—especially on the emergence and transformation of the State Department.

■ BARBARA ROSENBLUM (Style as Social Process) is Assistant Professor of Sociology at Stanford University. Her research centers on the structure of art reward systems. She is author of *Photographers at Work* (Holmes and Meier, 1978).

■ JAMES S. HOUSE (Comment on Hope, ASR June, 1975) is Associate Professor of Sociology at Duke University. His research focuses on occupational stress and health among blue-collar workers.

ERRATUM

An error occurred in the article entitled "Deviance and Moral Boundaries" (ASR August, 1976) by Pat Lauderdale. On page 667 (paragraph 2, line 14) the text reads, "The scale ranged from 'not at all' (position 1) to 'very much' (position 7)." The correct version is: "The scale ranged from 'not at all' (position 1) to 'very much' (position 9)."

New Publications Available from ASA...

_____ ***1978 Guide to Graduate Departments of Sociology***

Members & Students: \$4.00

Non-members & Institutions: \$10.00

_____ ***1978 ASA Directory of Members***

Members: \$5.00

Non-members & Institutions: \$10.00

_____ ***1978 Directory of Departments of Sociology***

Members: \$5.00

Non-members & Institutions: \$10.00

_____ ***Journal of Health & Social Behavior Index***

(Volumes 1-17, 1960-1976)

Members: \$5.00

Non-Members & Institutions: \$7.50

To order any of the above, check publication(s) desired and send coupon, enclosing payment, to: American Sociological Association, 1722 N Street, NW, Washington, DC 20036.

(All orders must be prepaid)

Name _____

Address _____

City

State

Zip

FACTORS AFFECTING RESPONSE RATES TO MAILED QUESTIONNAIRES: A QUANTITATIVE ANALYSIS OF THE PUBLISHED LITERATURE*

THOMAS A. HEBERLEIN

ROBERT BAUMGARTNER

University of Wisconsin, Madison

American Sociological Review 1978, Vol. 43 (August):447-462

Two hundred fourteen manipulations of the independent variables in 98 mailed questionnaire response rate experiments were treated as respondents to a survey, yielding a mean final response rate of 60.6% with slightly over two contacts. The number of contacts and the judged salience to the respondent were found to explain 51% of the variance in final response. Government organization sponsorship, the type of population, the length of the questionnaire, questions concerning other individuals, the use of a special class of mail or telephone on the third contact, and the use of metered or franked mail on the outer envelope affected final response independent of contacts and salience. A causal model of the final response rate, including initial response, explaining 90% of the variance, and a regression equation predicting final response rates are presented to show that high response rates are achievable by manipulating the costs of responding and the perceived importance of both the research and the individual response.

INTRODUCTION

Research on mail questionnaire response rates generally has studied the effect of a single factor, or a group of related factors, on the response rate while attempting to hold all other potential factors constant. Review articles by Scott (1961) and Linsky (1975) have presented published evidence on these factors affecting mail questionnaire response rates. In addition to these two general reviews, Armstrong (1975) reviews the evidence on

the effects of monetary incentives on response rate.

Ideally a very large factorial experiment, where factors potentially affecting mail questionnaire response rates are simultaneously varied, is needed. Even a cursory listing of these potential factors, however, suggests a factorial design of such proportions that one quickly appreciates a "second best" method. One such approach was introduced by Sudman and Bradburn (1974).

Using this procedure, we treated each study as a respondent in a survey. Data are coded from each report and quantitative comparisons are possible. Sudman and Bradburn (1974) used this method to examine response bias in interview surveys, and Glass and Smith (1976) have utilized a similar procedure to determine the effects of therapy on mental illness.

The approach has several limitations. Because the studies have been published, they are not as cooperative as survey respondents and often fail to provide infor-

* Support for this research was provided by the College of Agricultural and Life Sciences, University of Wisconsin, Madison, under Hatch Grant Project No. 5117. The authors would like to thank J. Stanley Black, Richard Campbell, Don Dillman, Kenneth Green, and an anonymous reviewer for this journal for their helpful comments on earlier drafts. The authors, however, take sole responsibility for the interpretations and conclusions in the article. This is a revised version of a paper presented at the annual meeting of the Rural Sociological Society, 1977, Madison.

mation about a specific variable. Thus missing data can become a problem. The factors or independent variables are correlated, unlike a true factorial experiment, and this can lead to difficulties in interpretation. Multivariate analysis can be helpful here to disentangle confounds, but the clarity of certain findings may be reduced.¹ Finally, it is sometimes difficult to establish a clear domain of studies to be included in an analysis.

In spite of these disadvantages, the approach does provide a clearer, more parsimonious review than previous qualitative discussions. The results are quantitative, so a researcher gets a sense of the possible impact of a procedure across all published studies, rather than an illustrative few selected by a reviewer.

In this research we coded 71 factors reported in 98 methodological studies of the mailed questionnaire and examined their effect on response rates. Methodological studies were selected since they tend to report sufficient information about procedures, which reduces the missing data problem. Moreover, an initial list of these studies was available. The primary goal of this research is descriptive: to identify the quantitative effect of a variety of factors on response rates.

The findings of our analysis give a sense of the actual process of mailed questionnaire research. The multivariate analysis shows which factors lower overall response directly or through the initial returns. We discovered what background characteristics lead a researcher to use multiple mailings and special contact procedures and the relative importance of these factors for increasing response rates. Finally, we try to interpret the results in a theoretical framework. The next steps, we hope, would be some experimental tests of our post hoc interpretations.

¹ Glass and Smith (1976) note that statistical adjustment by regression, which regresses the outcome variable in a study onto the descriptive features of a study, allows a researcher to avoid an arbitrary exclusion of certain studies with differing designs or variations in outcome measures. This method will enable the researcher to capture a greater share of the variation in both outcomes and situational variables of relevant studies.

METHODS

The primary source of the sample used in this study was Potter et al. (1972), a comprehensive bibliography of publications on the mail questionnaire. The bibliography contained 193 citations, 91 of which either were not experiments or did not report response rates (primarily validity or reliability studies), and 15 of which were textbooks. In addition, four of the citations could not be located. Of the remaining 83 citations which were mail questionnaire experiments involving response rates, 80 were coded since three of the citations were a further analysis of data already cited in the bibliography and coded for this study.

Eighteen additional response rate experiments were found by reviewing journals and citations in the published articles, giving a total of 98 independent experiments involving mail questionnaire response rates.² In many of the studies reported, the experimental design used more than one treatment of the independent variables and, in these cases, each treatment was coded separately. For the 98 studies coded, there were 214 treatments of the independent variables. Our unit of analysis was the individual treatment, since each treatment contained separate independent variables; so there are 214 cases. The number of treatments in each case was coded and retained for analysis.³ The sample includes a com-

² The citations for all of the studies coded as well as the correlation matrix for the exogenous variables in Figure 1, and six tables presenting the zero-order findings for all of the variables used in the analysis have been filed with the National Auxiliary Publications Service of the American Society for Information Science. See NAPS document # 03256 for twenty pages of supplementary material. Order from ASIS/NAPS, Microfiche Publications, P. O. Box 3513, Grand Central Station, New York, New York 10017. Remit in advance \$3.00 for Microfiche copy or, for photocopy, \$5.00 up to 20 pages. All orders must be prepaid. Institutions and organizations may order by purchase order. However, there is a billing and handling charge for this service. Foreign orders add \$3.00 for postage and handling.

³ A variable indicating the number of dependent treatments related to each observation was coded. For example, if a study reported five different manipulations of the independent variables, each was used as a separate case for analysis. The variable described above was then coded as five, indicating

prehensive set of mail questionnaire response rate research published as methodological articles. It does *not* include those articles published only as substantive research, using the mail questionnaire.

Variables

Seventy-one independent variables were initially coded. These fall into four categories: (1) general research characteristics; (2) sampling and sample characteristics; (3) questionnaire characteristics; and (4) research procedures.

General research characteristics include the year of publication, the journal of publication, the researcher's background (such as sociology, public health, or market research), the type of research organization (such as a university or a governmental agency); and the number of dependent treatments of the independent variables related to the particular individual treatment.

Sampling and sample characteristics included the sample size, the source of the sample (general, such as telephone directories or voter registration lists, organization records, etc.), the type of population sampled (general, nonprofessional occupations, students, etc.), the relation of the respondent to the study (no relation, employee, client, student, former student, subscriber or donor, etc.), and the selection of the respondent, indicating whether the questionnaire was addressed to a specific individual or to an unspecified person such as resident or head of household.

Variables measuring questionnaire characteristics include subject of the questionnaire (self, household or family, members of work organization, other individuals, respondent's work organization or place of residence) and type of data requested (attitude or evaluation, own behavior, other's behavior, ownership, and SES or demographic data). Each of these preceding variables was coded as a dichotomous dummy variable when reported in the studies coded.

that there were five dependent treatments of the independent variables related to each of these particular cases.

Another questionnaire characteristic was the saliency of the topic to the respondent. Saliency was rated by the second author on a three-point scale of salient, possibly salient or nonsalient. A salient topic was one which dealt with important behavior or interests that were also current. Examples of this category of salience are a Veterans Administration survey of the educational plans and interests of veterans, who had expressed an interest in the V.A. educational assistance programs, or a survey of Blue Cross plan enrollment directors and executive directors by the Health Information Foundation, a well-known information source in the field of public health.

Topics judged possibly salient were important issues or behaviors that were not necessarily current or timely. Examples are a survey of occupational status and mobility of former students of a university, or a survey of the political attitudes and characteristics of contributors to the National Committee for an Effective Congress.

Finally, topics judged to be nonsalient to the respondent were those that neither concerned important issues or behaviors nor were current, such as a survey of SES and demographic characteristics of readers of a national magazine or a study of consumer purchasing behavior in a geographical sample of urban households.

The length of the questionnaire was coded in three different ways: the number of questions, the number of pages in the instrument and the amount of time required to complete the questionnaire. A final questionnaire characteristic was the effect of the data on the respondent which was coded on a four-point scale of confidential, embarrassing, controversial or neutral.

The last category of variables, research procedures, included the total number of contacts, advance contact (dummy variable), total number of follow-up contacts, total number of waves (with replacement questionnaire), class of mail used for each contact (third or first class, special delivery or certified, telephone or personal contact), the length of time between each contact, use of incentives on each contact, use of a personally typed address (dummy

variable), and type of postage used on each envelope (metered or franked, stamp, multiple stamps).

The identifying procedures were coded on a three-point scale with the respondent's name on the questionnaire, the respondent's identification number only on the questionnaire, or the use of the anonymous return postcard technique.

Missing Data and Coding Procedures

Not all of the studies analyzed reported information for all of the independent variables. Where it was possible, assumptions were made about missing data to allow the construction of dummy variables. In some cases, unless the article explicitly stated that a procedure was used, we assumed it was not. This was done for three procedures: special delivery or certified mail, telephone or personal contact; the inclusion of monetary incentives; and the use of a personally typed address on the outer envelope.

For the three variables measuring length (the number of questions, the number of pages and the time to complete the questionnaire), information was reported for only 52, 86, and 18 of the 214 treatments, respectively. A letter was sent to those authors for whom addresses less than 10 years old could be located. Sixty-seven authors' addresses were located, representing 147 treatments, while 26 authors, representing 59 treatments, could not be located or were deceased. Forty-seven responses (70%) were obtained, yielding complete information on 90 of the 214 treatments of length. Combined with information from those studies whose authors had not been contacted or had not responded, we described 117 treatments in terms of length in number of questions, 126 treatments in terms of length in the number of pages and 90 treatments in terms of length in the time required to complete the questionnaire. The analysis of the effects of questionnaire length is based upon these sample sizes.⁴

⁴ An analysis of variance revealed some differences between the 126 studies reporting length, in pages, and the 88 studies which did not. Studies reporting length data tended to be more recent, have

Other variables with excessive missing data were omitted from the analysis, reducing the actual number of variables used in analysis to 55 from the 71 coded.

FINDINGS

On the average, 48% of those who received *one* mailing of a questionnaire returned it. The standard deviation is 19.9%. Some studies got better than an 80% response while others got less than a 20% response to one mailing, as is shown in Table 1. It seems unlikely that such tremendous variation is simply random, and the goal of our analysis will be to identify those factors which can explain at least part of this variability.

Contacts

Repeated mailings are widely used and have been reported by many researchers to increase returns (Lindsey, 1921; Rollins, 1940; Sletto, 1940; Ferriss, 1951; Goldstein and Kroll, 1957; Scott, 1961; Bachrack and Scoble, 1967; Champion and Sear, 1969; Dillman et al., 1974; Linsky, 1975). Table 1 shows that a follow-up mailing nets a return of nearly

a higher response rate and a higher number of total contacts, and used more follow-ups. The researchers were more likely to come from sociology and less likely to come from market research backgrounds, more likely to use a special class of mail on the third contact and anonymous post card procedures for coding returned questionnaires, and less likely to use postage stamps on the outer envelope containing the instrument. The studies were more likely to survey a sample of the general population, request data on the respondent's household or family, or his membership in a voluntary organization. They were less likely to be conducted by a private research organization, to sample nonprofessional and subscriber or donor populations, and to ask potentially embarrassing, confidential or controversial questions.

The three models were run with length, in pages, as an independent variable, for a complete data set where the mean number of pages of length was substituted for missing data, and for a reduced data set where using only the 126 observations reporting length. In spite of the differences between studies reporting length and those that did not, none of the regression coefficients were significantly different between analyses, nor were significant changes in R^2 observed.

From this we concluded that the lack of data on this variable does not seriously hamper conclusions, where length, in pages, is an independent variable.

Table 1. Response Rate Summary in Percents

Response from	n	Mean	SD	Percentiles				
				5	25	Median	75	95
Initial Mailing	183*	48.1	19.9	18.6	30.1	47.4	62.1	82.4
Follow-Up 1	58	19.9	7.7	7.5	14.0	19.5	25.6	31.6
Follow-Up 2	40	11.9	6.2	3.0	7.8	10.7	15.5	23.5
Follow-Up 3	25	10.0	5.1	2.8	6.8	8.1	14.3	19.3
Final	214	60.6	24.3	21.2	40.9	60.8	82.9	96.6

* 21 studies reported only a final response rate, after multiple contacts; hence, responses for initial mailings are available for only 183 treatments.

20% of the initial sample. A second and third follow-up yield about 12 and 10% returns, respectively. There is, again, substantial variability in the effectiveness of follow-ups, which makes simple generalizations difficult.

A follow-up may be considered as one of a variety of contacts between the researcher and respondent. Advance letters, postcards, letters, follow-ups which include an additional copy of the questionnaire, and even telephone calls are all examples of such contacts. Half the surveys analyzed ($N = 107$) had only one contact with the respondent. These showed a 46.1% response rate ($S.D. = 12.4$). Twenty-five studies had three contacts and responses went up to 80.6% for this group. Thirty-one had four contacts, but the response of 83.9 was not significantly better than for three contacts. There were too few cases to assess fully the effect of five and six contacts, but no noticeable increases over three contacts were observed. Contacts showed the highest zero-order correlation with response rates ($r = .634, p < .001$), accounting for 42% of the variance. In linear regression terms, each contact increases the predicted response rate by 12%.

Salience

Questionnaires also were more likely to be returned if they were judged to be salient to the respondent. Surveys with non-salient questionnaires averaged a 42% response ($N = 43$), while the 26 questionnaires judged to be salient for the respondent obtained a 77% return. The possibly salient group ($N = 112$) showed yields of 66%. This three-point salience scale is

correlated .427 ($p < .001$) with final response. Taken together, salience and contacts account for 50.5% of the variance in final response rates. The regression equation showing unstandardized coefficients for these two variables predicting the final response in 214 studies is:

$$\hat{Y} = 25.55 + 10.66 \text{ contacts} + 14.31 \text{ salience.}$$

Responses also varied with a number of other factors. For example, surveys published in market research journals showed only a 40% response, while studies published in scientific journals had a 65% response. Public health surveys showed an 81% response while university-based surveys showed 62% returns. *In most cases these differences are explained by the greater number of contacts and higher salience of the more effective surveys.* Consequently, to describe most parsimoniously the major results of this research, we focus on those variables which affect responses independently of contacts and salience. These findings are summarized in Table 2. None of the other independent variables measured had any effect net of contacts and salience.

General Research Characteristics

Government sponsored research got higher responses independent of contacts and salience. The unstandardized partial regression coefficient indicates that surveys under governmental sponsorship should expect an additional 12.4% responses than similar studies with equal numbers of contacts and similar salience to the respondent. Recent surveys get no

Table 2. Independent Variables Regressed on Final Response Rate Along with Contacts and Saliency (Unstandardized Coefficients)

	Coefficients for:				R ^a
	Constant	Contacts ^a	Saliency ^b	Individual Variables	
Base Equation	25.55 (2.70)	10.66 (.90)	14.34 (2.11)	—	.505
RESEARCH CHARACTERISTICS					
Government Agency Sponsor ^c	24.56 (2.66)	11.07 (.89)	13.05 (2.10)	12.38 (3.77)	.527
SAMPLE CHARACTERISTICS					
School and Army Sources ^c	25.27 (2.59)	9.57 (.90)	14.00 (2.02)	11.96 (2.75)	.544
Student and Army Populations ^c	25.43 (2.66)	10.00 (.92)	14.27 (2.08)	8.21 (3.04)	.520
General Populations ^c	28.58 (2.81)	10.92 (.88)	13.13 (2.10)	-8.06 (2.53)	.526
Employee Respondents ^c	24.91 (2.63)	10.97 (.88)	12.95 (2.08)	15.15 (4.17)	.533
Student Respondents ^c	25.15 (2.67)	9.80 (.94)	15.29 (2.11)	9.58 (3.54)	.520
QUESTIONNAIRE CHARACTERISTICS					
Length: Number of Questions ^a	28.88 (2.89)	10.95 (.89)	14.18 (2.07)	-.05 (.02)	.522
Subject: Other Individuals ^c	25.51 (2.68)	10.49 (.90)	13.92 (2.10)	9.76 (4.96)	.512
RESEARCH PROCEDURES					
Special Third Contact ^d	29.09 (2.88)	7.85 (1.27)	13.66 (2.08)	6.98 (2.26)	.525
Metered or Franked Postage ^c Outer Envelope	23.41 (2.79)	10.97 (.89)	14.67 (2.08)	8.91 (3.40)	.519

^a Variable coded as actual number (ratio variable).

^b Variable coded on a three-point scale.

^c Dummy variable.

^d Variable coded on a four-point scale.

Standard errors are in parentheses, all regression coefficients are at least twice as large as their standard errors.

All increases in R² over the base equation are statistically significant at the .05 level or beyond.

higher responses than questionnaire studies done in the 1940s and 1950s.

Sampling and Sample Characteristics

Students, employees, and military personnel are more likely to return questionnaires. Surveys of the general population are less likely to be returned than surveys of special subgroups. These zero-order relationships were reduced but remained statistically significant when controlled for contacts and saliency.

Length and Topic

A number of studies have examined length, finding either no effect or a modest

negative effect (Sletto, 1940; Scott, 1961; Mason et al., 1961; Champion and Sear, 1969; Berdie, 1973). In our data the average questionnaire had 72 questions on seven pages and took less than one-half hour to complete. For the variables measuring length in questions and number of pages, the standard deviations were larger than the means. Five percent of the questionnaires were over 22 pages long and had in excess of 230 items. In spite of this great variation in length, there was *no significant zero-order correlation between any of the length measures and overall responses*. Long questionnaires averaged just as high a response as very short instruments. When saliency and contacts

are controlled, instruments with more items get lower returns. As is shown in Table 2, each additional question reduces responses by .05%.

The content of the questionnaire appears to have little effect on response once saliency and contacts are controlled. Asking questions about SES, ownership, work organization, or the household, for example, did not lower or raise response rates. When the subject of the questionnaire was other individuals, respondents were *more* likely to return questionnaires, as is shown in Table 2.

Research Procedures

While the researcher cannot do much about many of the variables we have considered thus far, the actual research procedures such as certified mail, the number of contacts and incentives, are more susceptible to control. We have already shown that the number of contacts is very important. Although Linsky (1975) reviewed 12 studies where an advance contact was found to affect final response rates, we found that this zero-order effect vanishes when total number of contacts is controlled. Advanced contacts are no more or less effective than follow-up contacts, actually enclosing a new copy of a questionnaire with follow-up reminders did not appear to increase responses beyond the effect of the reminder itself. This variable did not survive controls for contacts, and comparing the 32 studies with two contacts showed a 62% response from the 16 with an enclosure and 65% from those without a replacement questionnaire ($t = .14$, NS). Use of a special mailing procedure, such as certified mail or special delivery, or a personal or telephone contact does continue to increase response rates after controls.

A number of dummy variables were created examining the type of postage used on mail-outs and returns, and no zero-order effects were observed. After controls, however, metered or franked postage on the outer envelope did have an effect.

Relatively few studies used incentives. Two studies paid \$1.00 and had 80% returns, nine paid 50¢ and had 66% returns,

and seven paid 25¢ and had a 45% return (all using only one contact). Most studies ($N = 187$) used no incentives, and these averaged a 62% response. There is some suggestion of a linear trend for incentives, but there was no significant zero-order effect.

A General Model Predicting Responses to Mailed Questionnaires

The three variable models described thus far have dealt with final response rates. A mailed survey, is, however, a process: a questionnaire is sent out, and the repeated mail or telephone contacts are made. Use of these procedures may depend upon initial response rates. The initial response can be considered causally prior to final responses, follow-up contacts, and the use of special procedures on follow-up contacts. All background factors such as type of sample, sponsorship of the project and so on are causally prior to these four variables. This allows the construction of a rough causal model.

Initial Response Rates

We began to develop a model by seeing how well an initial response or a response which did not involve more than one contact might be predicted. All variables which were causally prior to this response and which had a significant zero-order correlation with initial response were included in a regression equation. Two additional variables, the use of monetary incentives in the first mailing and the use of an advance contact were also included in the model in spite of nonsignificant zero-order correlations. Strong evidence both from behavioral psychology and economics would suggest incentives should have some kind of effect, and the general effectiveness of contacts would also lead one to be hopeful about the effect of advance contacts on initial response.

In accordance with standard theory-trimming procedures, nonsignificant variables were dropped from the analysis and the final model presented in Table 3 was computed. To determine if an important predictor variable had been omitted by

Table 3. Predicting Initial and Final Response Rates

Independent Variables	Standardized Coefficient	Unstandardized Coefficient	Significance Level
Initial Response ^a			
Constant	—	53.90	.000
General Population	-.305	-12.21	.000
Incentive-First Contact	.237	5.51	.000
Market Research Background	-.235	-10.69	.000
School or Army Sample	.217	9.37	.000
Saliency of Content	.181	5.95	.004
Employee Population	.181	12.02	.002
Subject-Voluntary Organization	-.139	-14.44	.01
Attitude Questions	-.129	-10.85	.03
Government Organization	.124	7.40	.03
Final Responses ^b			
Constant		6.82	.000
Initial Response Rate	.651	.85	.000
Number of Follow-Up Contacts	.555	10.90	.000
Special Third Contact	.126	4.11	.000
Length-Number of Pages	-.110	-.48	.000
Saliency of Content	.100	4.34	.000

^a $R^2 = .402$, $F_{9,204} = 16.89$ $p < .0000$.

^b $R^2 = .906$, $F_{5,200} = 413.47$ $p < .0000$.

our initial decision rule, all causally prior variables with zero correlations (those variables which were not included in the final equation) were entered *after* the nine significant variables predicting initial response rates. None of the variables showed significant partial correlations, or regression coefficients, with the nine variables in Table 3 controlled for, and the new coefficient of determination was not significantly different from the one reported in Table 3, which suggests that nothing important had been omitted from the model.

Perhaps the most interesting finding is that in spite of nine predictor variables, only 40% of the variance in initial responses could be explained. This contrasts with our earlier finding that 51% of the variance in final responses could be predicted from two variables, salience and contacts. This low predictability of initial response, no doubt, is due to the lack of variables assessing the design and layout of the instrument, the respondent's knowledge and affiliation with the signer of the cover letter, and the construction of the cover letter itself. These could not be coded from the published record, but no doubt affect initial responses.

Unstandardized regression coefficients

show the percentage change in response that a one unit increase in the independent variable is expected to yield. Four dummy variables each decrease response by better than ten percentage points. Lower initial response may be expected when the researcher has a background in the marketing area, when attitude questions are used, when a general population is surveyed, and when the subject of the questionnaire is voluntary organizations.

Moving one unit up a four-point incentive scale (where 0 = no incentive and 4 = \$1.00) increases response by 5.5%; likewise, moving one point up the salience scale increases responses 6%. If a sample is from a school or army source, a higher response can be expected, and if the research is done by a government organization, responses will also be higher (the latter two were coded as dummy variables).

Follow-Up Contacts

Variables causally prior to follow-up contacts which were used in the analysis for the initial and final response rates, including the initial response, were considered as potential predictors to the number of follow-ups used. Apparently, inves-

tigators do not react to low initial response rates by adding more follow-ups, since this variable showed no zero-order or partial relationship.

Studies of families or households, or those based on student populations tended to have one additional contact. As questionnaires get longer, investigators are more likely to use follow-ups. For each ten pages of questionnaire, .65 follow-ups were observed. Studies of ownership had fewer follow-ups, while questionnaires which asked questions about other individuals had more. Less than 30% of the variance in the number of follow-up contacts could be explained by the variables coded (see Figure 1).

Another procedure an investigator can adopt in the midst of the research process is the use of special procedures such as telephone calls or special mailing (certified or special delivery). The use of this technique had substantial correlations with final response when used on a second or third contact. For reasons of economy it seems most reasonable to wait until the third contact to use special contact procedures, since the number of nonrespondents is reduced. Hence, this procedure is entered into the model. Causally prior variables used in the analysis of initial and final responses were regressed on a three-point variable (telephone or personal contact = 3; special delivery or certified mail = 2; regular mail = 1). A special third contact is more likely to be used for surveys of the general population, with questionnaires representing data about other individuals, and when an advance contact is used. When a questionnaire contains questions asking for the respondent's attitude or evaluation, a special third contact is less likely to be used. Less than 20% of the variance in this variable could be explained by the coded variables (see Figure 1).

Final Response Rate

Table 3 also shows that a high initial response rate, the use of follow-up contacts, the use of a special third contact and a topic that is salient to the respondent all tend to produce a higher final response rate, while an increase in the length of a

questionnaire, measured in terms of the number of pages, tends to decrease the final response rate. These five variables account for over 90% of the variance in the final response rate.

Table 2 indicated that the sponsorship of a governmental organization (student, army, employee and general populations), questions about other individuals and the use of metered or postpaid mail had significant partial relationships with the final response rate when controlling for the number of contacts and the saliency of the topic to the respondent. Table 3 indicates, however, that none of these variables has a significant effect on the final response rate independent of the five variables predicting final response rates.

Table 3 also indicates that the length of a questionnaire, in number of pages, does not affect the final response rate until the effects of initial response, the number of follow-ups, the use of a special third contact and the saliency of the topic are controlled.⁵

The path model showing this process is presented in Figure 1. Only one variable, the saliency of the topic, that affects the initial response rate also affects the final response rate. The path model suggests that those factors which tend to decrease the initial response, such as surveys of the general population, market research questionnaires or questionnaires about membership in voluntary organizations, can be counteracted by using follow-up contacts and a special third contact.

Predicting Responses

As a researcher considers using a mailed questionnaire, there is always a question of what return can be expected. From the regression equation presented in Table 3, an estimate of final response can be obtained using the codes for the independent variables described in the text. Because of the great variance of response rates and the small sample size, the 95%

⁵ Table 2 showed that length, only when measured in number of questions, showed a significant partial relationship; however, when entered in the final response model, only length in number of pages showed a significant regression coefficient. The latter, then, was used in the final response model.

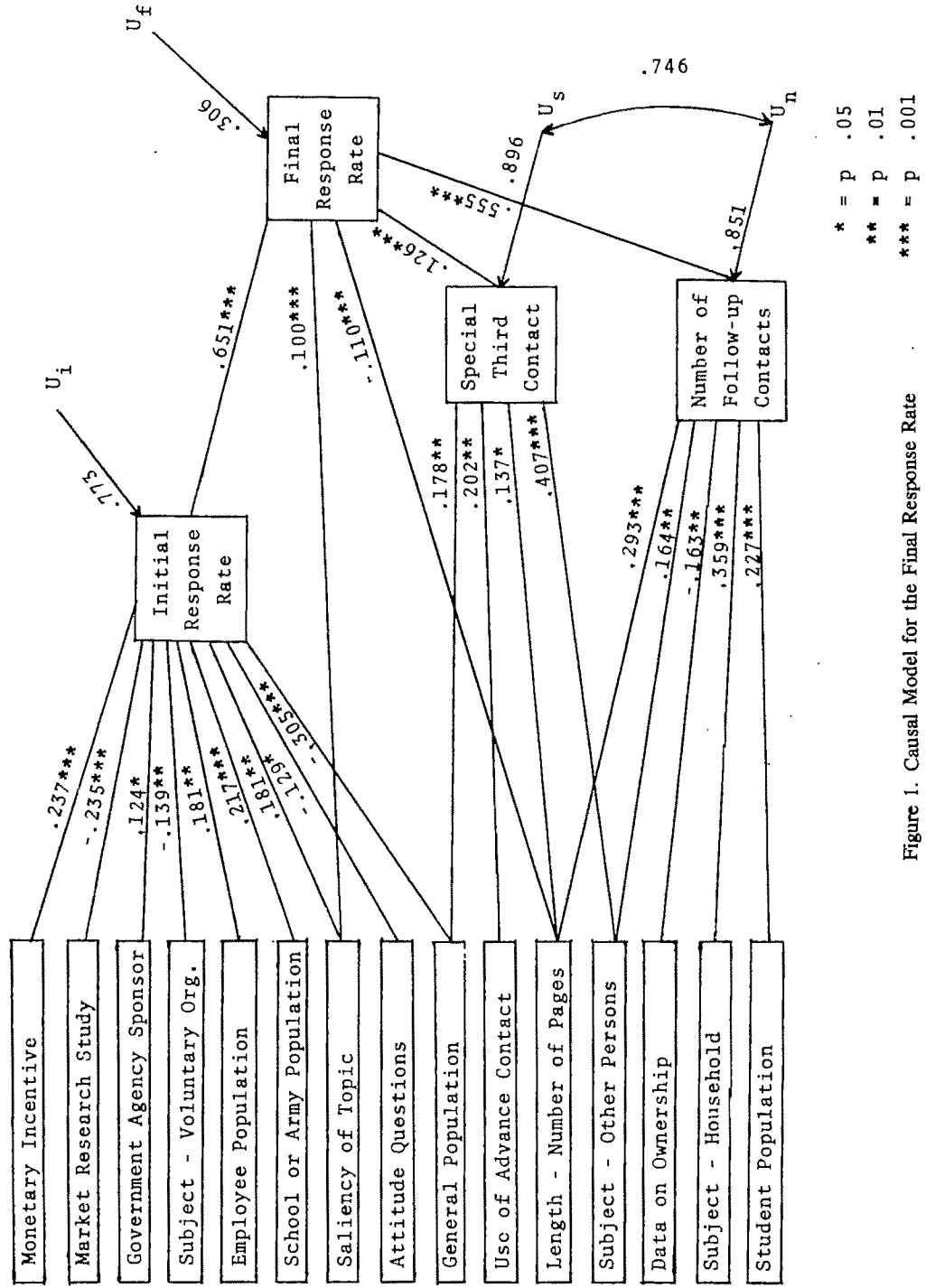


Figure 1. Causal Model for the Final Response Rate

confidence interval around a predicted response rate is $\pm 14.55\%$. The best information we can give an investigator, say with a prediction of 70% response, is that the observed response will lie between 55 and 85%, nineteen out of twenty times. An analysis of residuals shows, however, that for 66% of the 214 cases, predictions were within 5% of the estimate and 87% were within 10%. This may not seem very precise but it greatly improves on any prediction based on Table 1. From knowledge of the mean and standard deviation only, the 95% confidence interval would range from 14 to 108%; hence the multivariate model adds substantially to our ability to predict responses.

The more serious problem with the equation predicting final response in Table 3 is that to use it one needs an accurate estimate of the initial response. Usually one is interested in estimating final returns *before the study begins*. We have selected ten variables which are easily codable and which predict the final response rate when the initial response is *not* included in the equation. Using this equation in Table 4, we would predict that a market researcher surveying a general population on a topic of low salience with a four-page instru-

ment using only one contact, would have a projected response of $[36.3 + (-10.1) + (-7.5) + 0 + [4 \times (-.44)] + 7.4]$ or 24.3%. On the other hand, a 30-page government survey of a school population on a highly salient topic, using three contacts with regular mail on the third contact, would expect a response of 88.6% or $[36.3 + [30 \times (-.44)] + 10.2 + 9.9 + (2 \times 7.3) + (3 \times 7.4) + 8.6]$.

Unfortunately, the equation in Table 4 which explains 66% of the variance in final response, shows about twice as much variability, with a confidence interval of 28%. This means that 95% of the time a study with a projected response of 70% would yield an actual response between 42 and 98%. Analyzing the residual differences between projected and actual responses showed that the estimates are still useful. Thirty-one percent of the estimates were within 5% of the final response, 57% of the estimates were correct within a 10% margin, and 77% were within 15% of the predicted value. Using this model to predict responses to several unpublished surveys yielded predictions from 1 to 4% from actual returns.

If one predicts a low final response, new procedures can be introduced. Responses

Table 4. Ten Variable Model Predicting Final Response Rate

Independent Variable	Coefficient	Code Values
Constant	36.3	1
Market Research Background	-10.1	0-No, 1-Yes
Government Organization	10.2	0-No, 1-Yes
General Population	-7.5	0-No, 1-Yes
Employee Population	11.8	0-No, 1-Yes
School or Army Population	9.9	0-No, 1-Yes
Saliency of Topic	7.3	0 - Not salient 1 - Possibly salient 2 - Salient
Length-Number of Pages	-.44	Actual number
Total Number of Contacts	7.4	Actual number
Special Third Contact	8.6	0 - No third contact 1 - Regular mail 2 - Special mail 3 - Telephone or personal
Incentive-First Contact	6.1	0 - No incentive 1 - Less than \$.25 2 - \$.25 3 - \$.50 4 - \$1.00

$R^2 = .658$.

Standard error of estimate = 14.2.

can be increased most effectively by adding contacts, special contact procedures and incentives. Cutting the pages will help some but not as much as increases for the other variables. The predicted response of a market research sponsored, five-page survey of a general population would be 23.9%. By using three contacts, with a telephone call for the third contact, along with a \$1.00 incentive on the first mailing, the predicted response could be increased to 89%.

Besides illustrating the relative importance of variables, this prediction model shows several limitations of our analysis as well. The models are strictly linear. Our generalizations apply only for the observed range of variables. We don't suppose that ten contacts would really increase responses 74% over base. Non-linearity was not considered nor were all possible interactions examined. This may be particularly important with the incentive variable. Paying physicians \$1.00 to return a questionnaire may increase reactance and reduce response. The possible interactions need experimental attention.

DISCUSSION

Scott (1961) has pointed out that high response rates to mailed questionnaires are possible. Dillman and his associates (1974), writing more than a decade later, demonstrate that this is the case for general populations in the United States. Our data show that this is true, and generally has been the case for at least 35 years. Over a quarter of the studies reviewed obtained a final response of more than 80% using a mailed questionnaire. In spite of over 100 methodological studies assessing the mailed questionnaire during this period, response rates do not appear to have improved much.

Textbook authors (Selltitz et al., 1954; Goode and Hatt, 1952; Nachmias and Nachmias, 1976; and Kerlinger, 1965) indicate that low response rates are a defect of the mailed questionnaire. They also are correct. Fifty percent of the studies considered got less than 61% overall response, and one-fourth got less than 41%. These findings, based only on published

methodological studies reflecting editorial gatekeeping, are probably likely to reflect higher than usual responses.

The defect in the mailed questionnaire is not so much low response rates, as a great variability in response rates across investigators, subject populations, questionnaires, and procedures. The skilled researcher can successfully use the mailed questionnaire, but unlike the interview, which has the power of a face-to-face personal contact to stimulate response, the mailed questionnaire must rely on other techniques to assure response.

High return rates are due in part to lowering the costs involved in completing and returning a questionnaire. Hence investigators regularly include postpaid return envelopes to directly reduce costs for the respondent and design questionnaires to be as easy as possible to fill out. In addition to lowering costs, the investigator will attempt to increase the motivation of the respondent to overcome the cost barrier. It appears that the most effective way of doing this is to increase the perceived importance of the study and the individual's response to it.

Importance

When the content of the questionnaire is salient to the respondent and the respondent is knowledgeable and interested in the topic, the cost of responding may be reduced, and personal input to the study may be judged by the respondent as more important. Each additional contact further serves to convince potential respondents of the importance of their input (Robinson and Agisim, 1951). Contacts which show some special attention and greater expense and effort by the investigator also would seem to increase a sense of importance. Hence, special delivery, certified mail, telephone calls and personal contacts would be expected, as we found, to increase responses.

The auspices of the questionnaire can also affect the perceived importance. We found responses to government sponsored surveys to be higher. Information transmitted to the government may have a direct effect on the individual, or on other

similar individuals through policy changes. Similarly, the positive effects of metered or franked postage on the outer envelope may also be perceived as an indicator of importance. The use of this form of postage may be associated by respondents with governmental or other "official" research sponsorship.

Responses to other groups, such as scientists, may simply be less important to the respondent. Low importance would also explain why market research surveys net lower returns. Information about a purchase or one's demographic characteristics, used to benefit a firm, may not strike the respondent as particularly important to him- or herself.

Lower responses from general populations may also be explained in this framework. Specific appeals to self-interest or specialized questionnaire content are more difficult when sampling from driver's licenses or voter registration lists. It is still possible to achieve high response rates from general populations, but there is an initial inertia which must be overcome. Employees may be more likely to respond because their responses are seen as important to their livelihood.

The length is usually considered a cost barrier to be avoided if at all possible. Given our finding of no zero-order effects, but an observed negative effect of length only when other factors are controlled, it is possible that questionnaire length affects perceived importance in a way which tends to offset costs. Longer questionnaires may impress the potential respondent with the importance of his input. Tossing out a one-page instrument may be relatively easy to do, but discarding 30 pages of questions is depriving the investigator of a good deal of information. Also, if the researcher has taken the time to compose 30 pages of questions, it is clear that this research is a serious matter, not merely a passing curiosity. Length, then, may signal importance to the respondent, possibly even enough to overcome the costs associated with it.

Some investigators have found that personalization has had a modest effect on increasing response (Moore, 1941; Slocum et al., 1956; Linsky, 1965; Simon, 1967;

Carpenter, 1975). These increases could be viewed as an effect of importance, offsetting the reduction in importance when respondents view themselves as being anonymously selected on a random basis. Linsky (1975), however, discusses the published evidence on the use of different personalization techniques and reports almost as many studies reporting no advantage to personalization as those reporting an advantage. In our data we also find no general effect of personalization as measured by personally typed address or the use of postage stamps on mailing envelopes.

Costs

One way of reducing costs to respondents is to actually reimburse them for filling out a questionnaire. Armstrong (1975) has reviewed the published evidence on the use of monetary incentives and concludes that the larger the prepaid monetary incentive used, the greater the increment in the response rate. Our analysis suggests that the use of prepaid monetary incentives affects initial response and final response only when a number of other factors are controlled.

The costs of returning a mailed questionnaire may be lower for some groups, which in part can account for the observed higher returns. Students, for example, may find it easier to read and respond to questionnaires since they have been trained to carry out such cognitive tests, and these skills probably lower the cost barrier. Alternatively, or as a complement, both students and individuals obtained from military records, may find a request to complete a form more familiar and easier to comply with than other groups.

Longer questionnaires with more items clearly increase the cost factor and after a number of variables have been controlled, this cost barrier does exert a modest influence (about a 5% reduction in final response for every ten pages of questionnaire). Just as a longer questionnaire can possibly increase importance, it can also make the questionnaire *easier* to complete. Few things are worse than a variety

of questions compressed into one page. When respondents are asked to list *all* of their vacation expenditures, for example, on three one-inch lines, a questionnaire is difficult to complete even though it may be short. A listing of items, such as gasoline, motels, etc., followed by simple boxes which a respondent can check may be much easier to fill out, even though it takes six or ten pages to display. Long instruments tend to have fewer items per page, be less cluttered and reinforce the respondent's progress through page turning. All of these things tend to dissociate length from cost.

The negative effects of attitude questions on response may also be a function of costs to the respondent. Attitude questions often involve a response choice in which the individual may be ambivalent or undecided about the alternatives. Such cognitive exertion may be a sufficient cost to the respondents to deter some from completing the questionnaire.

Another potential cost for the respondent is the possibility that the information provided will invade privacy and implicate the individual. One way of reducing this cost without eliminating tracking so that follow-up contacts can be made efficiently is to adopt anonymous procedures, such as uncoded questionnaires and separate postcards which indicate that the instrument has been returned. Only seven studies reported using these techniques (Cahalan, 1951; Bradt, 1955; Boek and Lade, 1963; Vincent, 1964; Bachrack and Scoble, 1967; Dunning and Cahalan, 1973), but they obtained an 83% response compared to a 54% response for those studies with a code number on the instrument. Since this effect did not survive controls, it can only be interpreted most cautiously.

There are also certain costs in *not returning* a questionnaire, which may offset the costs of filling it out. The finding that employees are more likely to return questionnaires may reflect a concern for consequences revolving around a failure to comply with a request. The respondent may feel that potential occupational retaliations from not returning a questionnaire represent a higher cost than the time and effort to complete it.

Repeated mailings may also increase the cost of not responding. Follow-up mailings may serve to increase the guilt and lower the personal regard of individuals who intended, but "just never got around" to completing the instrument, independent of importance. Some of those who respond to repeated contacts may not feel that their reply is any more important, but that the psychological cost of not responding is simply greater than the time and effort to complete and return the questionnaire.

CONCLUSIONS

We proposed our methodology as a second best approach. The large factorial experiment still needs to be carried out. What our analysis does accomplish is to point to a more modest set of predictor variables which should be considered orthogonally. A factorial design could allow a better estimate of effects and interactions. The dependencies in the current data and the lack of cell size for more complex and interesting interactions limit our analysis.

Much of the mailed questionnaire research appears to be a post hoc tag-on to some other substantive research. What is not needed is another study reporting the effects of contacts, postage stamps or color of paper on a single instrument to a single population. It is time for the mailed questionnaire to become a research topic in its own right, much the way experimenter bias moved from a methodological to substantive matter in social psychology (Rosenthal and Rosnow, 1969; Rosenberg, 1969; Duncan et al., 1969).

Besides the orthogonal experiment which varies length, salience, incentives, types of populations and number of contacts simultaneously, it is time to look at the mediating factors. One needs to break into the questionnaire process and measure with the questionnaire itself, or through some other procedure such as telephone or personal interview, the respondent's reaction to systematically varied input such as cover letters and questionnaire format. We have theorized about such mediators, but they need empirical testing since we were not able to measure these variables, but discuss them largely

in a post hoc fashion to explain the observed effects on response.

More work needs to be done on *initial* rather than final response rates. Many studies have shown, and the current research has integrated and quantified, the effect of repeated contacts. What is largely unexplained are those factors which influence response to a single or initial mailing. Here also the social psychological reactions of respondents would be most important to determine.

Finally, mailed questionnaire research, totalling over 100 studies, is notably atheoretical. Most of the research rests more as demonstration projects than as a contribution to knowledge about human behavior. Yet, as an act, returning a questionnaire is not theoretically uninteresting. The return of a questionnaire represents a behavior on the part of the respondent and could well be a dependent variable for the important work being done in the attitude-behavior area (see Schuman and Johnson, 1976, for a review). Those theories useful for the prediction of behavior also should apply to filling out questionnaires, just as they apply to voting and other behaviors.

Altruism is another area with the potential of generating propositions which could be tested with responses to mailed questionnaires (Schwartz, 1977). Sending a person a questionnaire is a request for help. The circumstances under which individuals come to the assistance or fail to help another should apply to returning a questionnaire. Schwartz (forthcoming), for example, has found that women who did not return a questionnaire were less likely to volunteer for a helping task. He and Fleischman (1977) also found that those who returned questionnaires were more likely to volunteer than a sample who had not been given a questionnaire and consequently included both responders and nonresponders.

While space has prohibited any attempt, it seems possible to reinterpret our speculations about costs and importance within both the attitude-behavior and the helping-behavior theoretical frameworks. To do this would, we feel, be the most fruitful grounding for future mailed questionnaire research.

REFERENCES

- Armstrong, J. Scott
1975 "Monetary incentives in mail surveys." *Public Opinion Quarterly* 39:111-6.
- Bachrack, Stanley D. and Harry M. Scoble
1967 "Mail questionnaire efficiency: controlled reduction of non-response." *Public Opinion Quarterly* 31:264-71.
- Berdie, Douglas R.
1973 "Questionnaire length and response rate." *Journal of Applied Psychology* 58:278-80.
- Boek, Walter E. and James H. Lade
1963 "A test of the usefulness of the postcard technique in a mail questionnaire study." *Public Opinion Quarterly* 27:303-6.
- Bradt, Kenneth
1955 "The usefulness of a postcard technique in a mail questionnaire study." *Public Opinion Quarterly* 19:218-22.
- Cahalan, Donald
1951 "Effectiveness of a mail questionnaire technique in the army." *Public Opinion Quarterly* 15:575-8.
- Carpenter, Edwin H.
1975 "Personalizing mail surveys: a replication and reassessment." *Public Opinion Quarterly* 38:614.
- Champion, Dean J. and Alan M. Sear
1969 "Questionnaire response rates: a methodological analysis." *Social Forces* 47:335-9.
- Dillman, Don, Edwin Carpenter, James Christenson and Ralph Brooks
1974 "Increasing mail questionnaire response: a four state comparison." *American Sociological Review* 39:744-56.
- Duncan, S. J., Milton Rosenberg, and Jonathan Finkelstein
1969 "The paralanguage of experimenter bias." *Sociometry* 32:207-19.
- Dunning, Bruce and Don Cahalan
1973 "By mail vs. self-administered questionnaires." *Public Opinion Quarterly* 37:618-24.
- Ferriss, Abbott L.
1951 "A note on stimulating response to questionnaires." *American Sociological Review* 16:247-9.
- Glass, Gene V. and Mary Lee Smith
1976 "Meta-analysis of psychotherapy outcome studies." Unpublished paper, Laboratory of Educational Research, University of Colorado.
- Goldstein, Hyman and Bernard H. Kroll
1957 "Methods of increasing mail response." *Journal of Marketing* 22:55-7.
- Goode, W. J. and P. K. Hatt
1952 *Methods in Social Research*. New York: McGraw-Hill.
- Kerlinger, Fred
1965 *Foundations of Behavioral Research*. New York: Holt.
- Lindsey, E. E.
1921 "Questionnaires and follow-up letters." *Pedagogical Seminary Journal* 28:303-7.
- Linsky, Arnold S.
1965 "A factorial experiment in inducing re-

- sponses to a mail questionnaire." *Sociology and Social Research* 49:183-9.
- 1975 "Stimulating responses to mailed questionnaires: a review." *Public Opinion Quarterly* 38:82-101.
- Mason, Ward S., Robert J. Dressel, and Robert K. Bain
1961 "An experimental study of factors affecting response to a mail survey of beginning teachers." *Public Opinion Quarterly* 25:296-9.
- Moore, Clarence C.
1941 "Increasing returns from questionnaires." *Journal of Educational Research* 35:138-41.
- Nachmias, David and Chava Nachmias
1976 *Research Methods in the Social Sciences*. New York: St. Martin's.
- Potter, Dale R., Kathryn M. Sharpe, John C. Hendee and Roger N. Clark
1972 *Questionnaires for Research: An Annotated Bibliography on Design, Construction and Use*. USDA Forest Service Research Paper, PNW-140. Portland: Pacific Northwest Forest and Range Experiment Station.
- Robinson, R. A. and Philip Agisim
1951 "Making mail surveys more reliable." *Journal of Marketing* 15:415-24.
- Rollins, Malcolm
1940 "The practical use of repeated questionnaire waves." *Journal of Applied Psychology* 39:178-81.
- Rosenberg, Milton J.
1969 "The conditions and consequences of evaluation apprehension." Pp. 279-341 in R. Rosenthal and R. Rosnow (eds.), *Artifact in Behavioral Research*. New York: Academic Press.
- Rosenthal, Robert and Ralph L. Rosnow
1969 "The volunteer subject." Pp. 61-118 in R. Rosenthal and R. Rosnow (eds.), *Artifact in Behavioral Research*. New York: Academic Press.
- Schuman, Howard and Michael P. Johnson
1976 "Attitudes and behavior." Pp. 161-207 in Alex Inkeles (ed.), *Annual Review of Sociology*, Vol. 2. Palo Alto: Annual Reviews.
- Schwartz, Shalom H.
1977 "Normative influences on altruism." Pp. 222-79 in L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, Vol. 10. New York: Academic Press.
- Forth- "Temporal instability as a moderator of the coming attitude-behavior relationship." *Journal of Personality and Social Psychology*.
- Scott, Christopher
1961 "Research on mail surveys." *Journal of the Royal Statistical Society* 124:143-205.
- Selltiz, Claire, Marie Jahoda, Morton Deutsch and Stuart W. Cook
1959 *Research Methods in Social Relations*. New York: Holt.
- Simon, Raymond
1967 "Responses to personal and form letters in mail surveys." *Journal of Advertising Research* 7:28-30.
- Sletto, Raymond F.
1940 "Pretesting of questionnaires." *American Sociological Review* 5:193-200.
- Slocum, W. L., L. T. Empey and H. S. Swanson
1956 "Increasing response to questionnaires and structured interviews." *American Sociological Review* 21:221-5.
- Sudman, Seymour and Norman Bradburn
1974 *Response Effects in Surveys*. Chicago: Aldine.
- Vincent, Clark E.
1964 "Socioeconomic status and familiar variables in mail questionnaire responses." *American Journal of Sociology* 69:647-53.

MANUSCRIPTS FOR THE ASA ROSE SOCIOLOGY SERIES

Manuscripts (100 to 300 typed pages) are solicited for publication in the *ASA Arnold and Caroline Rose Monograph Series*. The Series welcomes a variety of types of sociological work—qualitative or quantitative empirical studies, and theoretical or methodological treatises. An author should submit three copies of a manuscript for consideration to the Series Editor, Professor Robin M. Williams, Jr., Department of Sociology, Cornell University, Ithaca, New York 14853.

A CROSS-CULTURAL STUDY OF EXPRESSIVE AND INSTRUMENTAL ROLE COMPLEMENTARITY IN THE FAMILY*

WILLIAM D. CRANO AND JOEL ARONOFF

Michigan State University

American Sociological Review 1978, Vol. 43 (August):463-471

Based on a representative cross-cultural sample of 186 societies, this research sought to extend an earlier investigation of specialization of the instrumental role, by sex, to a similar study of the expressive role in the family. A further aim of this research was the inspection of the Parsons and Bales proposition of complementarity of expressive and instrumental specialization. The results indicated that while women were strongly committed to expressive activities during the infancy of their children, the intensity of this commitment decreased substantially in early childhood. Males, while relatively less involved in expressive activities, still assumed a substantial portion of the expressive functions of the family, and maintained the level of this commitment throughout the infancy and early childhood of their offspring. A moderate level of male-female expressive complementarity was found in the infant caregiving activities; beyond infancy, however, the complementary relationship disappeared completely. Finally, no evidence was found to support a principle of complementary expressive-instrumental specialization in the family.

The principle of role complementarity has long served an important function in the sociological analysis of patterns of interdependence within the family (e.g., Durkheim, 1893; Parsons and Bales, 1955; Bell and Vogel, 1968). While the specific factors thought to affect intrafamilial relationships differed among theorists, most agreed that the family was an economic, as well as social unit, and thus specialized complementary economic (instrumental) and social (expressive) roles were expected to evolve to preserve the integrity and promote the efficiency of the unit. Recent cross-cultural research (e.g., Ember and Ember, 1971; Sanday, 1973; Aronoff and Crano, 1975) examining the degree to which sex-role specialization characterized the instrumental role has provided little support for this major premise of family organization. The present investigation extends this inquiry to a consideration of the degree of sex-related activity in the expressive role and, in combination with the findings from the earlier research series, examines the empirical status of a general principle of role complementarity in the family.

Of the various approaches taken to study the principle of role complementarity, that of Parsons and Bales (1955) was the most explicit in its postulation of the allocation rules that were thought to be involved. Based in large measure on Bales's systematic study of small group interaction, two distinct roles were identified—the task specialist, who was concerned primarily with instrumental operations, and the socioemotional specialist, who concentrated on the expressive interactions within the group. In generalizing from these empirical observations of the small group to the theoretical analysis of the family, a similar principle of role complementarity was assumed, with the allocation rules founded on biosocial assumptions of sex differences (e.g., the expressive psychological consequences that follow from the mother nursing the baby). Thus, the role allocation process characteristic of small groups would develop, in the family, as a function of sex differences, with males assigned the instrumental, and females the expressive roles: "[The] more instrumental role in the subsystem is taken by the husband, the more expressive the wife. . . . [The] husband has the primary adaptive responsibilities, relative to the outside situation . . . whereas the wife is primarily the giver of love" (Parsons and Bales, 1955:151).

* We would like to thank Lawrence A. Messé for his helpful review of an earlier draft of this paper, and an anonymous reviewer of our earlier research on this issue who suggested the possibility of this analysis.

This argument assumed that both instrumental and socioemotional specialist roles were necessary for the maintenance of the family group, that both roles could not be assumed by the same person at the same time, and that the mutually exclusive aspect of simultaneous assumption would result in the adoption of different characteristic role specialties by different individuals within the unit. Given the intense mother-child emotional interaction fostered in nursing, it seemed reasonable to hypothesize that the female naturally would assume the role of socioemotional specialist, and the male, virtually by default in this proposal, that of the instrumental specialist. In other words, a simple complementary theory of role differentiation by sex was proposed.

Not all, of course, found the rationale underlying this proposal convincing. Slater (1961), for example, would not accept as necessary the proposition that because expressive and instrumental roles could not be expressed simultaneously, they therefore *must* fall principally into the behavioral repertoires of different actors. Extending this reasoning to its ultimate, Slater (1961:297) argued that to maintain consistency one must agree that

If a special person is required to lead the laughing and playing . . . then it follows that still another person will be required to lead the weeping and mourning, since clearly a person cannot laugh and mourn at the same time.

Slater's critique pointed out simply that even if one accepts the proposition that expressive and instrumental behaviors cannot be expressed simultaneously by the same person, this does not logically preclude their expression by the same actor in different circumstances. Thus, the causal mechanism underlying the presumed male-instrumental, female-expressive role allocation system thought to characterize family interaction patterns was deprived of its logical grounding. At the empirical level, however, it was still tenable to maintain that males were primarily responsible for the instrumental, or subsistence-oriented activities of the family, that females were concerned principally with socioemotional maintenance,

and that these roles interacted in a complementary fashion.

Later empirical research has allowed for a data-based inspection of the validity of the hypothesized universal sex-role allocation rules, however, and in general, the results of these studies have not proved supportive of the Parsons-Bales approach. Addressing themselves to the first axiomatic proposition of Parsons and Bales's model, for example, several studies (e.g., Sanday, 1973; Aronoff and Crano, 1975) employed Murdock's (1967) comprehensive data archive to question the proposition of male primacy in the tasks of subsistence production (and, thus, the universal male assumption of the instrumental role). These studies have pointed out that while Murdock (1937; 1949), in an early statement of the proposition that Parsons and Bales greatly expanded, indeed had isolated single activities in which males or females engaged to the exclusion of the other sex, the critical question of the degree to which these activities jointly contributed to the common larder was left unanswered. Following procedures similar to those used by Brown (1963), Sanday (1973) and Nerlove (1974), Aronoff and Crano (1975) studied the degree to which the societies described by Murdock (1967) were dependent upon five central subsistence tasks (gathering, hunting, fishing, animal husbandry, and agriculture), and the degree to which males and females within each society contributed to each of these activities. This research procedure provided information of the relative contribution, by sex, to the subsistence input of the society, and thus allowed a more reasonable assessment of the proposition of universal male task specialization.¹ The results demonstrated that far from being a universal feature of the family, there was a wide range of variability of contribution by sex between societies, with females' contributions to their subsistence economies ranging from practically nothing in some groups to nearly 80% in others. Over all

¹ See Rosaldo and Lamphere (1974) and Aronoff and Crano (1975) for a discussion of the limitations of this operationalization of the instrumental role.

societies, females accounted for an average of nearly 44% of their societies' foodstuffs, with the range of contributions distributed such that in 45% of all societies surveyed, women were responsible for 40% or more of their group's subsistence production. It is simply not reasonable to maintain the principle of universal, or even predominant, male task specialization in light of results of this nature.

While Murdock's (1967) archive provided information enabling an assessment of the viability of the proposition of male primacy in the subsistence sector, it did not allow a test of the complementary principle of female specialization in the expressive role. What was needed was a source which indicated the distribution of expressive activities, by sex, across a representative range of societies. Later archival research by Barry and Paxson (1971) on a 186 society sample of Murdock's original population, however, provided data that permit an estimate of the nature of sex-related specialization on the expressive dimension.

SAMPLE, ANALYSIS AND RESULTS

Two years after the publication of the ethnographic atlas, Murdock and White (1969) selected a sample of societies from the atlas, chosen to minimize the problems of misrepresentation within subculture areas. As the total atlas attempted to catalogue information on all known societies, it was inevitable that some regions of the world would be represented more heavily than others. Accordingly, the total land area of the world was divided into 186 distinct culture regions, with a single representative society chosen from each in the compilation of a new, more representative sample of the world's societies. Information from Murdock's atlas, descriptive of a host of social, cultural, and economic variables was available for this new sample. Later reports (e.g., see Murdock and Morrow, 1970) added to the number of variables available in the new Standard Cross-Cultural Sample, as it was called, and of these, the codes produced by Barry and Paxson (1971) provide information directly useful

for the present research. In this last index, information is presented on a number of variables concerned with societal practices during infancy and early childhood. Not all of these variables are relevant to the question of the present investigation, and, unfortunately none allow for as precise an estimate as was possible in our assessment of females' proportional contributions to their subsistence economies. However, several of these codes supply information on the degree to which males and females attend to, and provide care for, infants and children. While these codes are not a measure of the emotional quality of the nurturance of the child, they do present the best indication presently available, for a large number of societies, of the level of socioemotional support provided to the children by various caretakers. Thus, Barry and Paxson's codes provide a plausible, if indirect, means to assess the distribution of child care within a society, and so permit a test of the proposition that socioemotional specialization is a universal feature of the family, and is assumed by the female member.

Caretakers and Companions

The most directly relevant codes in Barry and Paxson's source are contained in Column 13, which specify the relative importance of mothers among the child's caregivers, and Column 14, which specifies the relative importance of fathers in the caregiver role. The information contained in these codes provides the empirical basis for the analyses to be undertaken.² The classification of levels of parental involvement in these codes are as follows:

Column 13. (Mother's Involvement)

(a) Infancy

1. Almost exclusively the mother.
2. Principally mother, others have minor roles.

² It should be noted that Barry and Paxson's (1971) codes are based on the usual societal practices for male, middle-born children. This is an obvious limitation, but this archive is the best available source of information of its kind. As such, generalizations from our analysis should be undertaken with caution.

3. Principally mother, others have important roles.
4. Mother provides half or less of the care.
5. Mother's role is significant but less than all others combined.
6. Most care except nursing is by others.
7. Practically all care, including nursing, is by others.

(b) Early Childhood

1. Almost exclusively the mother.
2. Principally the mother, but others have important roles.
3. Child spends half or less of the time with mother.
4. Majority of time is spent away from mother.
5. Practically all the time is spent away from mother.

Column 14. (Father's Involvement)

The same categories are used for (a) infancy and (b) early childhood:

1. No close proximity.
2. Rare instances of close proximity.
3. Occasional or irregular close proximity.
4. Frequent close proximity.
5. Regular, close relationship or companionship.

Mother's role. In examining the degree of sex-specific investment in the socioemotional role from these data (see Tables 1 and 2), it is important to distinguish between the findings that obtain in infancy (to nine months) and those in early childhood (one year to four or five years of age). The results presented in the row totals indicate that, over all societies, the

mother is responsible to a major degree for the care of the infant. For example, in 71 (50%) of the 141 societies for which data were available, the mother was judged to assume the almost exclusive (13a-1), or the principal (13a-2), care of the infant, with other people taking only minor roles. In an additional 57 (40%) of the societies surveyed, infant care was viewed to be undertaken principally by the mother (13a-3), with others having important roles (these "important roles" assumed less than half of the caretaker responsibilities). In other words, in 90% of the sample the mother was the principal caretaker of the infant.

The pattern of mother's activity in caring for young children (Table 2), however, does not show a similar high level of involvement. In none of the societies was the young child's (as opposed to the infant's) care assumed "almost exclusively" by the mother, and in only 24% of the sample was the mother involved "principally" in the child's care (vs. 90% in the case of the infant). Further, in 32% of the sample, the "majority" or "practically all" of the young child's time was spent apart from the mother. While these results do not yet indicate who else might be involved in the care of the child, by themselves they do not provide strong support for a principle of universal concentration of mothers' involvement in the family's expressive activities. At best, the results of this analysis support only a modified principle of specialization in the earliest period of a child's life, and hardly appear to support a generalized express-

Table 1. Cross Tabulation of Mothers' and Fathers' Expressive Contributions in Infancy *

	Fathers' Contributions (from Col. 14a)					Total
	1	2	3	4	5	
Mothers' Contributions (Col. 13a)	None	Rare	Occasional	Frequent	Regular	
1. Almost exclusively	2	2	1	0	0	5
2. Principal (others minor roles)	4	14	32	16	0	66
3. Principal (others major roles)	1	8	24	22	2	57
4. 50% or less	0	0	6	3	1	10
5. Significant but less than all others	0	1	0	1	0	2
6. Most care by others	0	0	1	0	0	1
7. Practically all care by others	0	0	0	0	0	0
Total	7	25	64	42	3	141

* More complete specification of each of the categories included in these tables is presented in the text.

Table 2. Cross Tabulation of Mothers' and Fathers' Expressive Contributions in Early Childhood *

Mothers' Contributions (Col. 13b)	Fathers' Contributions (from Col. 14b)					Total
	1	2	3	4	5	
	None	Rare	Occasional	Frequent	Regular	
1. Almost exclusively	0	0	0	0	0	0
2. Principally	1	6	10	11	1	29
3. 50% or less	0	2	15	32	5	54
4. Majority of time with others	2	8	11	15	2	38
5. Practically all care by others	0	0	0	0	1	1
Total	3	16	36	58	9	122

* More complete specification of each of the categories included in these tables is presented in the text.

ive specialization in the female which is carried over into all areas of life, and make her the expressive "star" of the family.

Father's role. The contribution of the father to the caretaker activities of the children was characterized by a relatively lower level of expressive participation than that exhibited by the mother. Yet, from an absolute point of view, which is of great importance to the issue of sex-linked expressive specialization, an examination of the column total results presented in Tables 1 and 2 indicates that the father was involved to a significant degree in the expressive functions of the family. In 45 of the 141 societies surveyed (32%), for example, fathers maintained a "regular, close relationship" (14a-5), or "frequent, close proximity" (14a-4) with the infant. At the other extreme of this dimension, fathers in 33 (22%) of the total sample exhibited "no close proximity" (14a-1) or only "rare instances of close proximity" (14a-2), with the remainder of the fathers (46%) coded as having "occasional or irregular close proximity" (14a-3) to the infant.

The fathers' involvement with their young children shows an even higher degree of activity which, over all societies, approximates the level of the mother's involvement (compare the marginals in Table 2). Fathers in only 19 of 122 societies (16%) never, or only rarely were involved with the child; 30% were involved occasionally or irregularly with the child; and, most important, 55% of the sample maintained a regular or close relationship with the child. Although the somewhat different titles given to the cod-

ing categories by Barry and Paxson make exact comparison of mother's and father's involvement difficult, the overall pattern of results indicates a rough comparability of the degree of parents' efforts with their young children. Taken together, the findings from both the infancy and the early childhood analyses indicate a reasonably active paternal involvement in the caregiver role. At a minimum, these comparisons certainly dispute the proposition that the father's efforts are concentrated in instrumental activities outside the family.

Complementarity in the Expressive Role

A more direct test of the complementarity hypothesis requires an inspection of the joint activities of both parents. There are two forms in which this hypothesis may be phrased. In the least restrictive case, this hypothesis makes the prediction that an inverse relationship will exist between parents' levels of expressive activity. If one of the parents makes a considerable investment in expressive commitments, the other would be expected to devote proportionately less effort to such activities. The more restrictive statement of the complementarity hypothesis, as phrased by Parsons and Bales, above, would add to the first proposition the further qualification that the expressive role be assumed primarily by the mother. From the analyses presented above it is clear that the more restrictive complementarity hypothesis is not tenable since in both the infancy and the early childhood analyses fathers were involved with their children to a significant degree. Correla-

tions among the measures of expressive involvement, which provide an indication of the validity of the less restrictive complementary hypothesis, are presented in Table 3. The examination of these correlations shows that an interesting and substantially more complex pattern of relationship exists between these variables. If complementarity in the expressive role operated, we would expect to find strong negative correlations between fathers' and mothers' expressive involvements. This result was approximated by the moderately strong negative association between the parents' caretaker involvements with their infant ($r_{1,3} = -.29$, $df = 129$, $p < .001$). Simple inspection of the data presented in Table 1 indicates that, irrespective of sex, as the involvement of one parent with the infant increases, that of the other, to some extent, decreases. The pattern of results presented in Table 3, however, indicates that this restricted form of a complementary expressive relationship is limited narrowly to the practices that obtain with infants. Examination of the association of mothers' and fathers' involvement with young children shows a total lack of relationship ($r_{2,4} = .00$). This null correlation indicates that beyond infancy, whatever complementarity may exist in the parents' expressive involvement with their children totally disappears.

The information presented in the first two tables, along with two additional correlations from the matrix of Table 3, provide grounds for a possible explanation of this intriguing data pattern. The correlational analysis indicates that fathers' ex-

pressive commitments during the infancy and early childhood of their offspring were remarkably stable ($r_{3,4} = .76$, $df = 127$, $p < .001$). Apparently, fathers' prescribed expressive responsibilities are relatively clear-cut within societies, and do not appear to fluctuate over the course of their children's early development. A comparison of the marginals of Tables 1 and 2 concerned with fathers' commitments will serve to bolster this observation.

The correlations of mothers' expressive commitments during the infancy and early childhood of their offspring suggest a more variable role relationship. As presented in Table 3, the constancy of mothers' expressive commitment was not great ($r_{1,2} = .33$, $df = 120$, $p < .001$). This perhaps was to be expected, given the strong involvement of mothers and infants disclosed in the first table.

Taken together, these stability correlations suggest a possible explanation for the complementarity of expressive roles evidenced during infancy. We begin with the observation that infants are totally dependent on others for the satisfaction of their basic physiological needs. One of these basic needs is fulfilled by nursing. Yet, however noticeable an activity such as nursing may be, the care of an infant covers a much wider range of activities, and it well may be that it is in variation in these additional caregiving activities that the explanation of the complementary relationship may be found. The derivation of a causal statement from a set of correlations is always difficult, yet there is an intriguing explanation that fits this pattern of results which we wish to offer in a ten-

Table 3. Matrix of Correlations between Expressive and Instrumental Variables¹

	1	2	3	4	5
1. Mothers' Caretaker Involvement-Infant ^a	...	122	131	128	151
2. Mothers' Caretaker Involvement-Child ^a	.33*	...	115	112	126
3. Fathers' Caretaker Involvement-Infant	-.29*	.12	...	129	143
4. Fathers' Caretaker Involvement-Child	-.28*	.00	.76*	...	139
5. Females' Subsistence Contributions	-.02	-.10	.00	.02	...

¹ Because of missing data, correlations are based on varying *N*s. The specific *N* for any particular correlation is presented in the upper right triangle of this matrix. An analysis of the relationship between missing data on one variable and the information of the other indicators of this table disclosed no consistent pattern of results.

^a These variables have been reflected from Tables 1 and 2, so that higher scores indicate greater involvement in the activity in question.

* $p < .05$, two-tailed.

tative way. The stability of the fathers' involvement over the two time periods, as compared with the mothers', indicates that the degree of the fathers' activity may be based on some factor outside the present analysis. If this is true, then to preserve the life of the infant, other people—and here, most notably, one thinks of the mother—would be required to supplement a lessened contribution from the father. This additional female contribution thus creates a complementary male-female relationship in the infancy period. The early childhood period, however, includes a number of years—particularly the three to five year period—when the child in many societies requires much less supervision. This is also the period after the bulk of nursing has ended (average nursing duration in the Barry and Paxson sample is 32 months). With a lessened absolute need for care and a greater psychological and physical independence in activity, it is possible that the mother of young children is not so required to supplement a lesser paternal involvement. Thus, a complementary relationship, based on supplementary mothers' care, would not be required in this stage of life.³

Complementarity of Instrumental and Expressive Involvements

The most complete form of the role complementarity hypothesis concerns the degree to which complementarity exists *between* instrumental and expressive commitments. The data necessary for such an analysis consist of ratings of ex-

pressive involvements (as presented in the Barry-Paxson archive, and employed throughout this report) along with estimates of male and female contributions to the instrumental activities of their societies. If complementarity of expressive and instrumental behaviors exists, a negative relationship between male and female involvements in these two distinct roles would be expected.

Information on the extent of male and female contribution to the subsistence burdens of their societies was provided in Murdock's (1967) ethnographic atlas. As discussed earlier, Aronoff and Crano (1975) employed this archive to develop estimates of the relative importance of male and female subsistence contributions within societies, by determining the extent to which a society was dependent upon one of five subsistence traits, and the extent to which males and females contributed, independently, to each of these activities.⁴ The sum of these (dependency \times contribution) products provided a rough indication of male and female involvement in the instrumental activity of their societies. By correlating the estimate of females' subsistence contributions with the male and female expressive commitment indicators of Tables 1 and 2, we obtain a rough estimate of the extent of expressive-instrumental complementarity across all societies.⁵ The correlations between these variables are presented in the final row of Table 3. As is evident from these entries, there exists almost no support whatsoever for a general complementarity hypothesis. Consider, for example, those correlations involving women's subsistence contributions and the degree to which mothers assumed the role of caretaker of their infant ($r_{1,5} = -.02$) and young child ($r_{2,5} = -.10$). These correlations account for practically no variance. The extent to which mothers are involved in caretaker activities, that is, appears or-

³ The puzzling correlation between mother's involvement with the infant and father's involvement with the child ($r_{1,4} = -.28$, $df = 126$, $p < .001$) requires a word of comment. A positive relationship would have been understandable in that as the mother is concerned with the infant, the father would be supplementing her care with the older children. A negative correlation is difficult to explain meaningfully. A possible explanation for this effect might lie in the stability of the father's involvement over the two childhood periods, leading to identical correlations with mother-infant involvement. That is, since father's involvements in infancy and childhood are very similar, the correlation of both of these variables with another (e.g., mother's caretaker involvement in infancy) will result in two very similar correlation coefficients.

⁴ See Aronoff and Crano (1975:15-7) for an exact description of the operations employed in the development of subsistence estimates.

⁵ In that total male and female subsistence contributions were ipsative in Aronoff and Crano (1975), the use of either the male or the female subsistence estimates produces identical results in the correlation matrix.

thogonal to the extent of females' involvement in the subsistence activities of the society.

Correlations involving male involvement with expressive actions and females' subsistence contributions lend yet further support to this point, if any is needed. As presented in Table 3, it is clear that the extent of male expressive involvement with their infants had no relationship with females' (or males'—see fn. 5) instrumental contributions in either the infancy ($r_{3,5} = .00$) or the early childhood ($r_{4,5} = .02$) of their offspring.

The pattern of results supports Slater's critique of the proposed universal principle of role complementarity, in which he argued that Parsons and Bales's assumptions of a single dimension of specialization, bounded at one end by the expressive pole, and at the other by the instrumental, was unduly simplistic. At a minimum, two orthogonal dimensions representing greater or lesser instrumental involvement, and greater or lesser expressive involvement, seem more appropriate. The results of this study suggest that it is quite possible to have groups in which individual parents concentrated on different roles, on both roles to a moderate degree, or were involved to a major degree in both instrumental and expressive functions.

The sole exception to this general observation appears to occur in infancy. Here, it seems apparent that the precarious physiological status of the organism places certain minimum demands upon the caretaker. If these demands are not met, the infant perishes. Under these circumstances, some type of complementarity must evolve, if the minimal survival demands of the infant are beyond the physical or psychological resources of a single individual. The data of the present analysis support this reasoning, and suggest a degree of parental complementarity in caretaker responsibilities during the infancy of their offspring. While the strength of the correlational link between maternal and paternal expressive contributions is not really large ($r = -.29$), it is reasonably indicative of a limited, but genuine complementary relationship. Beyond these very limited confines, however, the evidence for a general principle

of within- or between-role complementarity is very weak. From earlier research (e.g., Aronoff and Crano, 1975; Sanday, 1973), for example, we find no support for an instrumental complementarity hypothesis. In addition, from the integration of these earlier instrumental indicators with the present expressive variables, it seems clear as well that there exists no compelling evidence of between-role complementarity.

As specialist roles do not appear to be dichotomous universally, it seems reasonable to suggest that further study of this problem abandon its past concern for a single independent explanatory mechanism based upon sex differences—whether purported or real. Rather, it may be more fruitful to conceptualize the study of role specialization in the family in terms of those variables that have been shown, more generally, to control the development of social differentiation.

REFERENCES

- Aronoff, J. and W. D. Crano
1975 "A re-examination of the cross-cultural principles of task segregation and sex role differentiation in the family." *American Sociological Review* 40:12-20.
- Barry, H., III, and L. M. Paxon
1971 "Infancy and early childhood: cross-cultural codes." *Ethnology* 10:466-508.
- Bell, N. W. and E. F. Vogel (eds.)
1968 *A Modern Introduction to the Family*. Glencoe: Free Press.
- Brown, J. K.
1963 "A cross-cultural study of female initiation rites." *American Anthropologist* 65:837-53.
- Durkheim, E.
[1893] *The Division of Labor in Society*. Trans. by
1933 George Simpson. Glencoe: Free Press.
- Ember, M. and C. R. Ember
1971 "The conditions favoring matrilineal versus patrilineal residence." *American Anthropologist* 73:571-94.
- Murdock, G. P.
1937 "Comparative data on division of labor by sex." *Social Forces* 15:551-3.
1949 *Social Structure*. New York: Macmillan.
1967 "Ethnographic atlas: a summary." *Ethnology* 6:109-236.
- Murdock, G. P., and D. O. Morrow
1970 "Subsistence economy and supportive practices: cross-cultural codes I." *Ethnology* 9:302-30.

- Murdock, G. P. and D. R. White
1969 "Standard cross-cultural sample." *Ethnology* 8:329-69.
- Nerlove, S. B.
1974 "Women's workload and infant feeding practices: a relationship with demographic implications." *Ethnology* 13:207-14.
- Parsons, T., and R. Bales
1955 *Family, Socialization and Interaction Process*. Glencoe: Free Press.
- Rosaldo, M. F., and L. Lamphere
1974 *Woman, Culture, and Society*. Stanford: Stanford University Press.
- Sanday, P. R.
1973 "Toward a theory of the status of women." *American Anthropologist* 75:1682-700.
- Slater, P.
1961 "Parental role differentiation." *American Journal of Sociology* 67:296-311.

SEX DIFFERENCES IN THE COMPLEXITY OF CHILDREN'S PLAY AND GAMES*

JANET LEVER

Northwestern University

American Sociological Review 1978, Vol. 43 (August):471-483

Play and games are situations in which important informal learning takes place. Specific attention is given to the social skills that emerge as a consequence of a particular style of play. The study, which draws upon a multiple-method design, reveals significant differences in the organization of play between the sexes. The primary difference concerns the complexity of the social setting in which play occurs. Boys' play is more complex than girls' play, as indexed by such attributes as role differentiation, interdependence between players, size of play group, explicitness of goals, number of rules, and team formation. The possible sources and consequences of this sex difference are explored.

The cognitive development theorists in psychology, most notably Jean Piaget, have traced the growth in knowledge and perceptions through the various stages of childhood. To date, little has been done to chart the parallel development of interpersonal skills needed as the child moves from the egocentric orientation of the family to the community of children found in the school. George Herbert Mead (1934) initiated this line of thought with his classic essay on the child's learning to regard the "self as object" and "take the role of the other." Unfortunately, few have followed Mead's example.

Significantly, both Mead and Piaget recognized the rich learning environment provided in play. Mead credits the child's

shift from aimless play to the realm of structured games as a crucial step in the development of role taking. Piaget (1965), through a close study of the game of marbles, meticulously explains how children develop moral values while they play rule-bounded games. Aside from Mead and Piaget, little attention has been paid to the world of play and games in the study of childhood socialization.¹

This study follows in the Mead and Piaget tradition by focusing on play and games as situations in which crucial learning takes place, but it goes beyond Mead's and Piaget's work in three important ways. First, Mead and Piaget each rests his analysis on a single game, whereas this study is based on a wide range of play and game activities. Second, both Mead and Piaget ignore sex differences in play. Mead's solitary example is the boys' game of baseball, but he does not tell us how

* This research was supported by an N.I.M.H. dissertation fellowship. I would like to thank Stanton Wheeler, Louis W. Goodman, and R. Stephen Warner for their guidance and criticism, and the Yale Child Study Center for its help in securing access to the public schools. I also thank John Meyer, Allan Schnaiberg, Marshall Shumsky, and the anonymous reviewers for their careful reading and discussion of an early draft of this paper.

¹ Among others who have recognized the importance of play in childhood socialization are Roberts and Sutton-Smith, 1962; Stone, 1971; Bruner et al., 1976.

girls, who are less familiar with team play, learn the same role-taking lessons. Piaget mentions, almost as an afterthought, that he did not find a single girls' game that has as elaborate an organization of rules as the boys' game of marbles, but he too fails to draw out the implications of his observation. A central concern of this study is to explore sex differences in the organization of children's play and to speculate on the sources as well as the potential effects of those differences.

Third, the paper highlights a specific dimension of play hitherto disregarded, namely, the *complexity* of the learning experience. I shall define complexity in more detail below, but it includes many of those attributes associated with the emergence of modern industrial society, such as division of labor, differentiation, heterogeneity, and rationalization (Simmel, 1955; Tonnies, 1955; Durkheim, 1893; Weber, 1967; Parsons and Smelser, 1956). My basic thesis is that the play activities of boys are more complex than those of girls, resulting in sex differences in the development of social skills potentially useful in childhood and later life.

METHODOLOGY

A variety of methods was used to gather as much data as possible in one year, 1972. In total, 181 fifth-grade children, aged 10 and 11, were studied. Half were from a suburban school and the other half from two city schools in Connecticut. The entire fifth grade of each school was included in the study. Three schools were selected whose student populations were predominantly white and middle-class—a choice made deliberately because of the possibility that race and class distinctions would confound the picture at this stage of exploratory research.

Four techniques of data collection were employed: observation of schoolyards, semistructured interviews, written questionnaires, and a diary record of leisure activities. The diary was a simple instrument used to document where the children had actually spent their time for the period of one week. Each morning, under the direction of the researcher, the children

filled out a short form on which they described (1) what they had done the previous day after school, (2) who they did it with, (3) where the activity took place, and (4) how long it had lasted. Half the diaries were collected in the winter and half in the spring. The questionnaire, designed to elicit how children spend their time away from school, also was administered by me inside the classroom. I conducted semistructured interviews with one-third of the sample. Some were done in order to help design the questionnaire and diary; others were done later to help interpret the results. I gathered observational data while watching children's play activity during recess, physical education classes, and after school.²

MEASURING COMPLEXITY

In common usage, the word "complex" means something that is made up of a combination of elements. Sociologists similarly have applied the term to describe the amount of functional differentiation in any social unit, from a small group or a large organization, to society as a whole. Based on the ideal type of complex organization, regardless of the scale of the collectivity, there is general agreement that increases in any of the following six attributes constitute greater complexity (Etzioni, 1969; Blau and Schoenherr, 1971):

1. division of labor based on specialization of roles;
2. interdependence between individual members;
3. size of the membership;
4. explicitness of the group goals;
5. number and specificity of impersonal rules; and
6. action of members as a unified collective.

Borrowing from the work of some contemporary students of games (Roberts et al., 1959; Redl et al., 1971; Avedon, 1971; Eifermann, 1972), I developed operational definitions for these six dimensions of complexity as they apply to the structure of play and games:

² See Lever (1974:65-108) for a detailed description of the methodology.

1. *Role differentiation.* For the purposes of this study, activities are to be considered low in role differentiation if the same behavior is required or expected from all players. For example, in the game of checkers, each player is equipped with the same number of pieces and is expected to move them in accordance with the same rules. Role differentiation is to be scored medium if one player has more power and acts differently from the undifferentiated group of other players. This describes all central-person games such as tag and hide-and-seek. An activity is to be scored high on role differentiation if three or more distinct game roles are present. For example, in the game of baseball, the pitcher has a different task to perform than the shortstop whose task is different from the center fielder and so on.

2. *Player interdependence.* An activity is to be judged low on the dimension of interdependence of players when the performance of one player does not immediately and significantly affect the performance of other players. For example, in the game of darts, one person's score does not interfere with the next player's score for the round. On the other hand, in the game of tennis, each player's move greatly affects the other's so that game has high interdependence of players.

3. *Size of play group.* This is a simple count of the number of players engaged in an activity. In this analysis, a group of three or fewer children is considered low on this dimension of complexity.

4. *Explicitness of goals.* The explicitness of goals is found in the distinction between play and games. *Play* is defined as a cooperative interaction that has no stated goal, no end point, and no winners; formal *games*, in contrast, are competitive interactions, aimed at achieving a recognized goal (e.g., touchdown; checkmate). Goals may involve tests of physical or mental skills, or both. Formal games have a predetermined end point (e.g., when one opponent reaches a specified number of points; end of ninth inning) that is simultaneous with the declaration of a winner or winners. The same basic activity may be either play or games. For example, riding bikes is play; racing bikes is a game.

5. *Number and specificity of rules.* Sometimes the word "rule" is broadly used to refer to norms or customs. Here the term is used in a narrower sense and refers to explicit rules which (a) are known to all players before the game begins, (b) are constant from one game situation to the next, and (c) carry sanctions for their violation. Play as defined above never has rules, whereas games always are governed by them. But games do vary by the number and specificity of their rules. Some games, like tag and hide-and-seek, have only a few rules; other games, like baseball and monopoly, have numerous well-established rules.

6. *Team formation.* A team is a group of players working collectively toward a common goal. Play, as defined above, is never structured by teams. Games, on the other hand, are to be divided into those requiring team formation when played with three or more persons and those prohibiting or excluding team formation. Within the category of games with team formation are included both those games where teammates play relatively undifferentiated roles, as in tug-of-war or relay races, and those that require coordination between teammates playing differentiated positions, as in baseball.

In order to test the hypothesis that boys' play and games are more complex in structure, I examine closely the type and frequency of the play activities of both sexes as they occur in public and private places. The evidence for private play is in the diary data, reporting after school and weekend play. Diary data are important because they reflect a large number of incidents, a wide range of activities, and a free choice of both games and playmates. The evidence for public play, based on observational data collected mostly during recess and gym periods, reveal the rich texture of the play world, replete with dialogue that helps the researcher understand the meanings children attribute to different play forms.

Diary Data

The diary responses reflect activities played inside or around the home in the

hours after school. From over two thousand diary entries, 895 cases of social play were isolated for this analysis.³ They represented 136 distinct play activities which were then scored by the author and three independent coders.⁴ The operational definitions of the six dimensions of complexity were presented to the coders, along with descriptions of play activities derived from the children's interviews. The activities were then rated along each of the six dimensions. All games were given ratings based on the children's own

reports of how a game is played most typically at the fifth grade level.⁵

Table 1 presents the basic data. To develop an overall complexity score, five of the dimensions were dichotomized and assigned either a low or high value (0,1).⁶ (The sixth dimension, size of group, varied from one play situation to the next and was tabulated independently.) The five dichotomous attributes yield thirty-two possible combinations; however, only nine occurred empirically. In Table 1 they are organized from lowest to highest complexity (scores from 0 to 5). Only the forty-eight activities that appeared in the diaries ten or more times are used to exemplify this scoring procedure, but all social play activities, even those less frequently mentioned, are included in the tabulations.⁷ By age ten, play activities

³ There were 2,141 activities recorded in the children's diaries. Five hundred eight entries were eliminated from this analysis because they were descriptions of nonplay activities like attending church services, doing homework or household chores, or going to the doctor. Another 527 items were eliminated because they reflected pastimes rather than actual play. This category included: watching television; reading books, comics or newspapers; going to the movies; going for an auto ride; and talking on the telephone. Television viewing, by far, accounts for most of the entries in this category. Of the remaining 1,106 play activities, 211 were not included because they were instances of the child's playing alone rather than in the company of others. Because the complexity dimensions reflect interpersonal skills, pastimes and solitary play are not relevant. However, it should be noted that there was no sex difference in the number of leisure hours spent with the television (15 to 20 hours/week) or playing alone (about 20% of all play).

⁴ The coders included the headmistress of a private elementary school who previously had taught fifth-graders for over a decade, a graduate student who had been a camp counselor for ten-year-olds for several years, and an assistant professor of sociology. Overall, the judges agreed on over 90% of the items coded.

⁵ Such reports were especially needed because separate groups of children may play the same game somewhat differently, while even the same children do not necessarily play a game in identical fashion from one occasion to the next. It is also important to note that children modify adult games, so that a game like pool, which has complicated rules for adults, usually is played according to simple rules by children.

⁶ To justify linking the six dimensions, a factor analysis was run on the 136 activities. There was only one factor present, and all six dimensions were a part of it (the lowest degree of communality was .60); I have referred to this single factor as "complexity." While it may be argued that some dimensions add more complexity than others, the absence of guidelines encourages equal weighting at this time.

⁷ See Lever (1974:394-7) for a complete list of games recorded in the children's diaries.

Table 1. Coding and Complexity Scores of the Most Frequently Listed Diary Activities

	Girls	Boys	Total
Type I. Complexity Score=0			
one role (0); low interdependence (0); play (0)			
no rules (0); no teams (0).			
1. listen to records (g)			
2. listen to radio (g)			
3. drawing (g)			
4. painting (g)			
5. work with clay (g)			
6. build things (b)			
7. ice skating (g)			
8. roller skating (g)			
9. bike riding			
10. mini-biking (b)			
11. exploring woods (b)			
12. hiking (g)			
13. horseback riding (g)			
14. grooming horses (g)			
15. take a walk (g)			
16. jump roofs (b)			
17. climb trees (b)			
18. sled ride (b)			
19. launch rockets (b)			
20. fly kites (b)			
Type I:	42%	27%	34%
	(179)	(126)	(305)

are generally known to be sex segregated. The "g" or "b" after each activity in Table 1 indicates whether it is played predominantly by girls or boys; the absence of a letter implies that the sexes engage in the activity with roughly equal frequency.

Table 1 yields two important findings. First, it shows the great variety regarding levels of complexity in the games played by children of similar age. Mead and Piaget, by focusing on only a single game, could not show the range of experiences available within the play world. Fully a third of the activities were low on all the measured dimensions of complexity. Another fifth were high on all. Children exposed to one or the other of these types of play are likely to be learning very different skills. Second, if we can agree that games provide differential learning environments, then we must assume differential effects for boys and girls. Boys experience three times as many games at the highest level of complexity and over twice as many boys' activities are located in the top half of the complexity scale.⁸

Table 2 views the data from a different perspective by showing the sex distribution separately for each of the six dimensions. Although greater complexity in boys' activities is demonstrated for all six, the major finding of Table 2 is seen on the fourth dimension, explicitness of goals. Sixty-five percent of boys' activities were competitive games compared to only 37% of girls' activities. In other words, *girls played more while boys gamed more*. This difference is not merely a function of boys' playing more team sports. Only 140

Table 2. Sex Differences on the Six Dimensions of Complexity in Play and Games

Dimensions of complexity	Girls	Boys
1. Number of roles (3 or more roles)	18% (427)	32% (468)
2. Interdependence of players (high interdependence)	46% (427)	57% (468)
3. Size of play group (4 or more persons)	35% (427)	45% (468)
4. Explicitness of goals (game structure)	37% (427)	65% (468)
5. Number of rules (many rules)	19% (427)	45% (468)
6. Team formation (teams required)	12% (427)	31% (468)

of the 305 games played by boys were team sports. Eliminating team sports for both sexes, we would still find 54% of the boys' activities and 30% of the girls' activities competitively structured. Sedentary games, like chess and electric race cars, are as important as sport in reflecting boys' greater competitiveness.

Nor is it the case when girls do participate in competitive games that they experience the same level of complexity as their male peers. The games girls play have fewer rules, and less often require the formation of teams.⁹ In summary, the data from children's diaries show strongly that boys, far more often than girls, experience high levels of complexity in their play and games.

Observational Data

Observations of children at play during recess, gym classes, and after school also indicate very distinct play patterns for boys and girls. As in the diary data, boys' activities were found to be more complex. The following descriptions of a few selected play activities illustrate the way in which each of the dimensions of complexity is expressed. Greater attention is given to girls' games as they are less familiar to adults. Some implications of differ-

⁸ Because some children reported more activities than others, there is the possibility that these results, based on activities as the units of analysis, reflect the extreme scores of a few individuals and are not representative of the sample as a whole. To guard against such misinterpretation, I made the individuals the units of analysis. To do so, I used the same dichotomization and point system displayed in Table 1 and added the sixth dimension, size of play group, as it appeared in each of the 895 entries. Once each activity could be given a complexity score (now zero to six), an average complexity score could be ascertained for each child based on the entire week's social play report. Seventy percent of the boys, compared with 36% of the girls, had average complexity scores of 3.0 or higher—a fact which further sustains the hypothesis.

⁹ Fifty-one percent of the games girls play ($n = 158$) contain many rules, compared with 69% of the boys' games ($n = 305$). Looking only at games with three or more participants, we note that boys played 26% more games which called for team formation.

ential organization of play are suggested, but their elaboration awaits the discussion section.

1. *Role differentiation.* The largest category of girls' public activity was the same as their private activity, namely, single-role play. These were cooperative activities with both or all parties doing basically the same thing such as riding bikes, roller skating, or ice skating. A minority of girls' activities were competitive games. Observing recess periods for a year, I saw only one instance of a spontaneously organized team sport, namely, kickball. The activities that appeared most regularly during recess were the traditional girls' games, like hopscotch, which are turn-taking games with only one game role present at a given time. Each player, in specified sequential order, attempts to accomplish the same task as all other players. A few turn-taking games have two distinct roles: for example, in jump-rope there is the role of rope turner and that of rope jumper. The other girls' games I observed frequently at recess were central-person games, the most popular being tag, spud, and Mother May I. These games also have only two roles—the "it" and the "others." Power is usually ascribed in these games through "dipping rules" like "odd-man-out."

Boys at this age have largely stopped playing central-person games except as fillers; for example, they might play tag while waiting for a bus or after so many team members have been called home to dinner that their previous game has disintegrated. The great majority of observed games were team sports with their multiple roles. Besides distinctions based on positions and assigned tasks, there were also distinctions in power between team captains and their subordinates. Sometimes the leaders were appointed by teachers, but more often the children elected their captains according to achievement criteria.¹⁰ After school espe-

cially, I observed boys in single role activities, some noncompetitive, like flying kites and climbing trees, but most competitive like tennis, foot races, or one-on-one basketball.

2. *Player interdependence.* There are many types of player interdependence: (1) interdependence of action between members of a single group; (2) interdependent decision making between single opponents; (3) simultaneous interdependence of action with one's own teammates and an opposing group of teammates.

Very little interdependence was required of those girls engaged in single role play; coaction rather than interaction is required of the participants. Also, little interdependence was required of those playing turn-taking games. Even though the latter activity is competitive, the style of competition is indirect, with each player acting independently of the others. That is, one competes against a figurative "scoreboard" (Player A → norm ← Player B). Participation in such games is routinized and occurs successively or after the previous player's failure; that is, opponents do not compete simultaneously. Interdependent decision making is not necessary in turn-taking games of physical skill as it may be in some of the popular board games.

When girls do play interdependently, they tend to do so in a cooperative context where there is interdependence of action between members of a single group. This type of interaction is best exemplified (but rarely observed) in the creation of private fantasy scenarios. One public example occurred when seven girls from one school took the initiative to write, produce, and act out a play they called "Hippie Cinderella." They stayed indoors at recess and rehearsed almost daily for three weeks in preparation for presentation to the entire fifth-grade class.

When boys compete as individuals, they are more likely to be engaged in direct, face-to-face confrontations (Player

¹⁰ In response to the interview question "Who are the fifth-grade leaders?" the boys in all three schools answered that the best athletes/team organizers rightly held that position. In contrast, most girls hesitated with the question, then named persons who had power, but credited their aggression rather than particular valued skills. They equated giving directives

with assertiveness and gave that behavior negative labels like "bossy" or "big mouth." Some openly stated that leaders acted less than ladylike and were not envied for their power. Attitudes that underlie Kanter's (1977:201) "mean and bossy woman boss" stereotype obviously are set at a very young age.

A ↔ Player B). Interdependent decision making between single opponents is necessary in games like tennis or one-on-one basketball that combine strategy with physical skill. More often, boys compete as members of teams and must simultaneously coordinate their actions with those of their teammates while taking into account the action and strategies of their opponents. Boys interviewed expressed finding gratification in acting as representatives of a collectivity; the approval or disapproval of one's teammates accentuates the importance of contributing to a group victory.

3. *Size of play group.* Observations made during recess periods showed boys playing in much larger groups than girls to a far greater extent than appeared in the diary data. Boys typically were involved in team sports which require a large number of participants for proper play. Boys in all three schools could play daily, depending on the season, in ongoing basketball, football, or baseball games involving ten to twenty-five or more persons. Girls were rarely observed playing in groups as large as ten persons; on those occasions, they were engaged in cooperative circle songs that seemed to emerge spontaneously, grow, and almost as quickly disintegrate. More often, girls participated in activities like tag, hopscotch, or jumprope, which can be played properly with as few as two or three participants and seldom involve more than five or six. In fact, too many players are considered to detract rather than enhance the fun because it means fewer turns, with longer waits between turns. Indeed, Eifermann (1968), after cataloging over 2,000 children's games, observed that most girls' games, like hopscotch and jacks, can be played alone, whereas the great majority of boys' games need two or more players.

4. *Explicitness of goals.* In the recess yards, I more often saw girls playing cooperatively and boys playing competitively. Some girls engaged in conversation more than they did in play (see Lever, 1976:481). Others, like those who initiated the circle songs and dances, preferred action governed by ritual rather than rules.

For example, the largest and most enthusiastic group of girls witnessed during the year of research was involved in a circle chant called "Dr. Knickerbocker Number Nine." Twenty-four girls repeated the chant and body motions in an outer circle, while one girl in the center spun around with eyes closed. She then stopped, with arm extended, pointing out someone from the outer circle to join her. The ritual chant began again while the new arrival spun around; this procedure continued until nine persons had been chosen in similar random fashion to form the inner circle. Then the ninth person remained in the circle's center while the others resumed their original positions and the cycle would begin anew.

Although this activity appeared monotonous to the observer because it allowed the participants little chance to exercise physical or mental skills, these ten-year-olds were clearly enjoying themselves. Shouts of glee were heard from the circle's center when a friend had been chosen to join them. Indeed, a girl could gauge her popularity by the loudness of these shouts. For some the activity may provide an opportunity to reaffirm self-esteem without suffering any of the achievement pressures of team sports.

Even when girls engaged in presumably competitive games, they typically avoided setting precise goals. In two schools, I observed girls playing "Under the Moon," a popular form of jumprope. The first person hops in and jumps once, in any fashion of her choosing, and then hops out. She then enters again and does two jumps, usually though not necessarily, different from the first. She increases her jumps by an increment of one until she has jumped ten times. Her turn over, she then becomes a rope turner. There was no competition exhibited between players. They participated for the fun of the turn, not to win. Even if the jumper trips the rope, she is allowed to complete her turn. If the jumper competes, it is with herself, as she alone determines whether to attempt an easy jump or a more difficult one.

The point is that girls sometimes take activities in which a comparison of rela-

tive achievement is structurally possible (and sometimes normatively expected) and transform them into noncompetitive play. Girls are satisfied to keep their play loosely structured. For example, in the game of jacks, girls can say before beginning, "The first to finish 'double bounces' is the winner." More often, however, they just play until they are bored with the game. Players may or may not verbalize "you won," and recognize who has advanced the most number of steps. Boys grant much more importance to being proclaimed the winner; they virtually always structure their games, be it one-on-one or full team basketball, so that the outcomes will be clear and definite.

5. *Number and specificity of rules.* This investigator also observed, reminiscent of Piaget, that boys' games more often have an elaborate organization of rules. Girls' turn-taking games progress in identical order from one situation to the next; prescriptions are minimal, dictating what must be done in order to advance. Given the structure of these games, disputes are not likely to occur. "Hogging" is impossible when participation is determined by turn-taking; nor can fouls occur when competition is indirect. Sports games, on the other hand, are governed by a broad set of rules covering a wide variety of situations, some common and others rare. Areas of ambiguity which demand rule elaboration and adjudication are built into these games. Kohlberg (1964) refines Piaget's thesis by arguing that children learn the greatest respect for rules when they can be used to reduce dissonance in ambiguous situations.¹¹

Because girls play cooperatively more than competitively, they have less experience with rules per se, so we should expect them to have a lesser consciousness of rules than boys. On one of those rare occasions when boys and girls could be watched playing the same games, there was striking evidence for a sex difference in rule sensitivity. A gym teacher introduced a game called "newcombe," a simplified variation of volleyball, in which

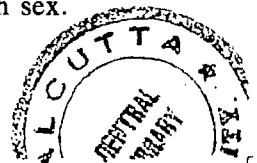
the principal rule is that the ball must be passed three times before being returned to the other side of the net. Although the game was new to all, the boys did not once forget the "3-pass" rule; the girls forgot it on over half the volleys.

6. *Team formation.* Team formation can be seen as a dimension of complexity because it indicates simultaneously structured relationships of cooperation and conflict. In turn-taking games, girls compete within a single group as independent players, each one against all others. Boys compete between groups, acting interdependently as members of a team. Team formation is required in all of their favorite sports: baseball, football, basketball, hockey, and soccer. Only a few girls in each school regularly joined the boys in their team sports; conversely, only a few boys in each school avoided the sports games. Questionnaire data support these observations. Most boys reported regular participation in neighborhood sports games. In addition, at the time of the study 68% said they belonged to some adult-supervised teams, with a full schedule of practice and league games. In fact, some of these fifth graders were already involved in interstate competitions.

The after-school sports program illustrates boys' greater commitment to team competition. Twenty girls from the third, fourth, and fifth grades elected captains who chose teams for newcombe games. Only seven of those girls returned the following week. In contrast, after-school basketball attracted so many boys that the fifth graders were given their own day. The teacher called roll for the next two weeks and noted that every boy had returned to play again.

Thus observational data, like the diary data, support the basic hypothesis that boys' play activities are more complex in structure than those of girls. Boys' play more frequently involves specialization of roles, interdependence of players, explicit group goals, and larger group membership, numerous rules, and team divisions. This conclusion holds for activities in public as well as in private. It suggests a markedly different set of socialization experiences for members of each sex.

¹¹ See Lever (1976:482-3) for a description of sex differences in the handling of quarrelling in games.



DISCUSSION

Sources of the Sex Difference

What is it that produces these distinct play patterns for boys and girls? The answer is mostly historical and cultural and holds true for much of Europe as well as the United States. While the rise of recreational physical activities in the late nineteenth century was enjoyed by women and men alike, the organized team sports which flourished at the same time were limited to participation by males (Paxson, 1917). The combined beliefs in the masculine nature of sport and the physiological inferiority of females led early twentieth century educators to lobby for competitive athletics for boys while restricting the physical education of girls to gymnastic exercises and dance. The emphasis on competitive athletics for males was reinforced by the view that sport served as a training ground for future soldiers ("the battle of Waterloo was won on the playing fields of Eton") and by the growing interest in spectator sports in which the dominant performers were young men (Cozens and Stumpf, 1953). Despite some outstanding individual female athletes in golf, tennis, and track and field, there was no development of interest in women's team sports. This situation is only now beginning to change.

Evidence generated in connection with Title IX of the Education Amendments Acts shows the extraordinary sex difference with respect to the allocation of funds for athletic programs from the primary grades through college. In 1969 the Syracuse New York School board allocated \$90,000 for boys' extracurricular sports compared to \$200 for girls' sports. In rural Pennsylvania, the Fairfield area school district set its 1972-73 budget at a ratio of 40:1 in favor of male athletes whose interscholastic competition begins in earnest by fifth grade. Even at Vassar, where sports for women are given great attention, the boys' athletic budget was double that of girls, although they comprised only one-third of the student body (Gilbert and Williamson, 1973).

Of course, it is not only the schools that encourage boys' and restrict girls' athletic participation. Parents act as the conveyor

belts for cultural norms, and it is no less the case for norms pertaining to sport. Male children are quick to learn that their demonstrations of athletic skill earn the attention and praise of adults. Many fathers show more emotion and enthusiasm for professional sports than anything else. Girls at young ages may not be actively discouraged from sports participation, but they are told that they are "tomboys" which is understood to be a deviant label. In the recent Little League debate, psychologists, parents, and coaches voiced their concern for the masculinization of female athletes, and the possible damage to young male egos when girls defeat boys in public (Michener, 1976). This cultural legacy is still with us, even though we now appear to be on the verge of radical change.

Historical analysis of children's games confirms that boys are playing more team sports now than ever before. Equally important, boys have drifted away from loosely structured play towards more formally organized competitive games (Sutton-Smith and Rosenberg, 1971). Evidence presented here supports this picture. It appears that the growing cultural emphasis on sports and winning has carried over to nonphysical activities and made them more competitive, and that, to date, it has had this effect to a far greater extent for boys than for girls.

Consequences of the Sex Differences

Boys' games provide a valuable learning environment. It is reasonable to expect that the following social skills will be cultivated on the playground: the ability to deal with diversity in memberships where each person is performing a special task; the ability to coordinate actions and maintain cohesiveness among group members; the ability to cope with a set of impersonal rules; and the ability to work for collective as well as personal goals.

Team sports furnish the most frequent opportunity to sharpen these social skills. One could elaborate on the lessons learned. The rule structure encourages strategic thinking. Team sports also imply experience with clear-cut leadership positions, usually based on universalistic

criteria. The group rewards the individual who has improved valued skills, a practice which further enhances a sense of confidence based on achievement. Furthermore, through team sports as well as individual matches, boys learn to deal with interpersonal competition in a forthright manner. Boys experience face-to-face confrontations—often opposing a close friend—and must learn to depersonalize the attack. They must practice self-control and sportsmanship; in fact, some of the boys in this study described the greatest lesson in team sports as learning to “keep your cool.”

Girls' play and games are very different. They are mostly spontaneous, imaginative, and free of structure or rules. Turn-taking activities like jump rope may be played without setting explicit goals. Girls have far less experience with interpersonal competition. The style of their competition is indirect, rather than face to face, individual rather than team affiliated. Leadership roles are either missing or randomly filled.

Perhaps more important, girls' play occurs in small groups. These girls report preferring the company of a single best friend to a group of four or more.¹² Often girls mimic primary human relationships instead of playing formal games, or they engage in conversation rather than play anything at all. In either case, there are probable benefits for their affective and verbal development. In Meadian terms, it may be that boys develop the ability to take the role of the *generalized other* while girls develop empathy skills to take the role of the *particular other*.

That the sexes develop different social skills in childhood due to their play patterns is logical conjecture; that those social skills might carry over and influence their adult behavior is pure speculation. Indeed, the weight of evidence indicates that life experiences are vast and varied;

much can happen to intervene and change the patterns set during childhood. Still, there is so much continuity between boys' play patterns and adult male roles that we must consider whether games serve a particular socializing function.

This idea is now popular. In a recent best seller on managerial leaders, MacCoby (1976) describes the 250 executives he studied as gamesmen who organize teams, look for a challenge, and play to win. The same social skills may be equally helpful in lower level bureaucratic jobs or other settings, like trade unions and work crews, where complexity of organization is also found. One need not endorse the world of organizations, bureaucracy, sharp competition, and hierarchy to recognize it as an integral part of modern industrial society.

The unfortunate fact is that we do not know what effect playing games might have on later life. We do not know, for example, whether the minority of women who have succeeded in bureaucratic settings are more likely to have played complex games. A recent study offers a modicum of supporting data. Hennig and Jar-dim (1977) portray their small sample of twenty-five women in top management positions as former tomboys. It is also the case that elite boarding schools and women's colleges, many of which stress team sports, have been credited with producing a large portion of this nation's female leaders. I would not want to argue that competitive team sports are the only place to learn useful organizational skills. Surely, the skills in question can be learned in nonplay settings in both childhood and adulthood. Nevertheless, it can be argued that complex games are an early and effective training ground from which girls traditionally have been excluded.

CONCLUSION

Children's socialization is assumed to have consequences for their later lives. Sociologists have looked to the family and the school as the primary socializing agents. In contrast, this analysis focused on the peer group as the agent of socialization, children's play as the activity of socialization, and social skills as the prod-

¹² It is important to note that, according to their questionnaire responses, the minority of thirty girls who reported playing complex games during the diary week also indicated a preference for larger friendship groups. The fact that the sex difference in size of friendship cliques disappears when controlling for complexity of game experience is one indication of the importance of this classification scheme.

uct of socialization. The data presented here reaffirm Mead's and Piaget's message that during play children develop numerous social skills that enable them to enjoy group membership in a community of peers.

The data also demonstrate that some games, when analyzed structurally, provide a highly complex experience for their young players while others do not. By itself, the notion of complexity adds to our appreciation of games as important early training grounds. However, the evidence of differential exposure to complex games leads to the conclusion that not all children will learn the same lessons. Here the approach to play and games differs dramatically from that of Mead and Piaget who presumed social and moral development as a normal part of the growth process and, therefore, did not make problematic the different experience of boys and girls. One implication of this research is that boys' greater exposure to complex games may give them an advantage in occupational milieus that share structural features with those games. At the very least, the striking similarity between the formula for success in team sports and in modern organizations should encourage researchers to give serious attention to play patterns and their consequences.

REFERENCES

- Avedon, Elliott M.
1971 "The structure elements of games." Pp. 419-26 in Elliott M. Avedon and Brian Sutton-Smith (eds.), *The Study of Games*. New York: Wiley.
- Blau, Peter and R. A. Schoenherr
1971 *The Structure of Organizations*. New York: Basic Books.
- Bruner, J. S., A. Jolly, and K. Sylva
1976 *Play: Its Role in Development and Evolution*. New York: Penguin.
- Cozens, Frederick and Florence Stumpf
1953 *Sports in American Life*. Chicago: University of Chicago Press.
- Durkheim, Emile
[1893] *The Division of Labor in Society*. New York: Free Press.
- Eifermann, Rivka
1968 "School children's games." Final Report, Contract No. OE-6-21-010. Department of Health, Education and Welfare; Office of Education, Bureau of Research. Unpublished paper.
- 1972 "Free social play: a guide to directed playing." Unpublished paper.
- Etzioni, Amitai
1969 *A Sociological Reader on Complex Organization*. New York: Holt.
- Gilbert, Bil and Nancy Williamson
1973 "Sport is unfair to women." *Sports Illustrated* 38 (May 28):88-98.
- Hennig, Margaret and Ann Jardim
1977 *The Managerial Woman*. New York: Doubleday.
- Kanter, Rosabeth Moss
1977 *Men and Women of the Corporation*. New York: Basic Books.
- Kohlberg, Lawrence
1964 "Development of moral character and moral ideology." Pp. 383-431 in M. L. Hoffman and L. W. Hoffman (eds.), *Review of Child Development Research*, Vol. 1. New York: Russell Sage.
- Lever, Janet
1974 *Games Children Play: Sex Differences and the Development of Role Skills*. Ph.D. dissertation, Department of Sociology, Yale University.
- 1976 "Sex differences in the games children play." *Social Problems* 23:478-87.
- Maccoby, Michael
1976 *The Gamesman*. New York: Simon and Schuster.
- Mead, George Herbert
1934 *Mind, Self, and Society*. Chicago: University of Chicago Press.
- Michener, James A.
1976 *Sports in America*. New York: Random House.
- Parsons, Talcott and Neil J. Smelser
1956 *Economy and Society: A Study in the Integration of Economic and Social Theory*. London: Routledge.
- Paxson, Frederic L.
1917 "The rise of sport." *Mississippi Valley Historical Review* 4:144-68.
- Piaget, Jean
1965 *The Moral Judgment of the Child*. New York: Free Press.
- Redl, F., P. Gump, and B. Sutton-Smith
1971 "The dimensions of games." Pp. 408-18 in Elliott M. Avedon and Brian Sutton-Smith (eds.), *The Study of Games*. New York: Wiley.
- Roberts, John M., M. J. Arth, and R. R. Bush
1959 "Games in culture." *American Anthropologist* 61:597-605.
- Roberts, John M. and Brian Sutton-Smith
1962 "Child training and game involvement." *Ethnology* 1:166-85.
- Simmel, Georg
1955 *The Web of Group Affiliations*. Trans. by Reinhard Bendix. New York: Free Press.
- Stone, Gregory P.
1971 "The play of little children." Pp. 4-17 in R. E. Herron and Brian Sutton-Smith (eds.), *Child's Play*. New York: Wiley.
- Sutton-Smith, B. and B. Rosenberg
1971 "Sixty years of historical change in the game preference of American children."

- Pp. 18-50 in R. E. Herron and B. Sutton-Smith (eds.), *Child's Play*. New York: Wiley.
- Tonnies, Ferdinand
1955 *Community and Association*. Trans. by Charles P. Loomis. London: Routledge.
- Weber, Max
1967 *From Max Weber: Essays in Sociology*. Trans. and ed. by H. H. Gerth and C. Wright Mills. New York: Oxford University Press.

THE TRANSITION TO ADULthood: SEX DIFFERENCES IN EDUCATIONAL ATTAINMENT AND AGE AT MARRIAGE*

MARGARET MOONEY MARINI

Battelle Human Affairs Research Centers

American Sociological Review 1978, Vol. 43 (August):483-507

The transition to adulthood involves a number of role changes, but the timing of the transition in large part, is determined causally by educational attainment and age at marriage. These two variables are related positively for both sexes, but because women marry earlier than men, the relationship is stronger for women. Analyzing data from a fifteen year follow-up study of the high school students studied by Coleman in The Adolescent Society (1961), this paper examines sex differences in the determinants of educational attainment and age at first marriage and in the relationship between these two variables. The results, based on estimation of a simultaneous-equation model, indicate that women's earlier age at marriage is an important factor limiting their educational attainment. Age at first marriage has no significant effect on the educational attainments of men but has a stronger effect on the educational attainments of women than any of the variables usually considered in male models of the educational attainment process. Educational attainment also exerts a stronger effect on the timing of marriage for women than men, although it has a significant effect for both sexes.

In the United States no one event marks the transition to adulthood. Beginning in late adolescence, when school attendance is no longer mandatory, a number of role changes begin to take place which mark the passage to adulthood. These include movement out of the student role, entry into the labor force or housewife role, entry into marriage, and entry into parenthood. Although these events are usually closely spaced in time, the point in an individual's life at which they begin to occur is influenced strongly by the duration of time spent in school and, hence, the level of educational attainment. Some

individuals begin to take on adult roles shortly after leaving high school in the late teenage years. Others who continue their educations beyond high school begin to take on adult roles later, usually after college or other additional schooling has been completed.

Entry into one's first full-time (non-summer) job tends to follow directly upon first leaving full-time schooling (Ornstein, 1975:28, 133; Marini, 1977b). The ages at which these two events occur are therefore highly related and both are a function primarily of the amount of schooling attained (Ornstein, 1975:26-8; Marini, 1977b).¹ Variables such as socioeconomic

* This is a revised version of a paper presented at the Population Association of America Annual Meeting, Atlanta, 1978. This research was supported by grants from the Russell Sage Foundation and the National Institute of Child Health and Human Development (1-R01-HD-09637-01). I am indebted to Peter H. Rossi and James S. Coleman for their assistance and support during the initial stages of this research. The assistance of Carin S. Weiss and Elizabeth J. Thomson in processing the data and Susan M. Ball in typing the paper is also gratefully acknowledged, as are the comments of the anonymous ASR reviewers.

¹ If one chooses to define entry into the labor force as the first full-time job after the completion of full-time schooling, age at last leaving full-time education and age at entry into the labor force are, by definition, also highly related and are a function of the amount of schooling attained. Evidence of these relationships previously has been confounded by problems of measurement (Blau and Duncan, 1967:166-7; Duncan et al., 1972:210-6), but more recent investigations are attempting to solve these problems (Featherman and Carter, 1976).

background, intelligence, and high school grades, which affect the amount of education attained (Alexander and Eckland, 1974; Sewell and Hauser, 1975:91-105), are therefore also likely to be the major causal factors affecting the age at leaving full-time education and age at entry into the labor force.

Age at first marriage and age at first birth, on the other hand, although related to the amount of education attained (Voss, 1977; Waite and Moore, 1978), are less directly a function of it (Marini, 1977b). To some extent, these family role transitions occur independently of the economic role transition marked by leaving school and entering the labor force. Age at first marriage and age at first birth themselves are related highly, however. Age at first marriage largely determines age at first birth since it usually sets a lower limit on the age at which first birth occurs. Only about 15% of births are premarital or result from premarital conception (Ryder and Westoff, 1971:299). Thus, although the transition to adulthood involves a number of role changes, the timing of the transition, in large part, is causally determined by educational attainment and age at marriage.

That educational attainment and age at marriage are related but also somewhat independent factors determining the timing of the transition to adulthood is evident from changes which have occurred in these variables over the course of the twentieth century. On the one hand, the level of education has increased, pushing back the age at first leaving full-time education and the age at entry into the labor force (Coleman et al., 1974:25-7; U.S. Bureau of the Census, 1976a). On the other hand, the age at marriage has declined, primarily as a result of factors such as the substantial postponement of marriage during the depression of the 1930s, earlier biological maturation, a rise in income, and an increase in the size of the marriage market as a result of the declining importance of ascribed relative to achieved characteristics in mate selection (Coleman et al., 1974:59-60; U.S. Bureau of the Census, 1976b).² Even the increase

in years of schooling may have led to a decline in age at marriage because educational institutions serve the latent function of a marriage market, bringing together large numbers of students of each sex, who are residentially or socially selected for homogeneity, on an age-graded basis (Coleman et al., 1974:60-1).

As the level of educational attainment has increased and the age at marriage has declined, the relationship between educational attainment and age at marriage at the individual level also is likely to have become more pronounced. When the amount of education attained was such that the normal age at leaving school was considerably below the age at marriage, educational attainment was largely irrelevant to the timing of marriage and vice versa. As the age at marriage has declined, however, differences in educational attainment among individuals are likely to have become more strongly linked to differences in age at marriage. The overall effect of these changes has been to create a relationship between the level of education (and, hence, the level of socioeconomic status) an individual attains and the timing of the major role transitions that mark the passage to adulthood.

Sex Differences in the Transition to Adulthood

If the relationship between educational attainment and age at marriage has become more pronounced in recent decades, it has become particularly pronounced for women. Because of the strain toward consistent differences between marriage partners across a variety of status dimensions (sex, age, education, occupation, earnings, height, physical strength), and because the traits which have been valued traditionally in each sex are differentially associated with age, it has long been common practice for women to marry men who are older than they are and, therefore, for women to marry at younger ages than men. Because of their younger age at marriage, women's educational and

reversal of this trend, and age at first marriage appears to be increasing (Coleman et al., 1974:59-62).

² Since 1960, however, there has been a slight

marital experiences are more likely to directly impinge on one another, making the relationship between age at marriage and educational attainment stronger for women than men. At present, women marry, on the average, about two years earlier than men (U.S. Bureau of the Census, 1976b) and, therefore, also have their first child about two years earlier. Women also receive less education than men (U.S. Bureau of the Census, 1976a) and are therefore likely to leave school at a somewhat earlier age. To the extent that they enter the labor force immediately after the completion of schooling, they are likely to begin their first full-time job somewhat earlier as well.

What remains unknown about the earlier transition to adulthood by women than men is the extent to which women's earlier age at marriage is responsible for cutting short their educational attainments. Specifically, it is not known to what extent age at marriage influences educational attainment and, conversely, to what extent educational attainment influences age at marriage. On the one hand, the amount of education an individual of either sex decides to achieve is likely to influence when she or he is free and either ready or financially able to marry. Since women acquire less education than men, they are free to marry earlier. On the other hand, women's earlier age at marriage actually may cause them to terminate their schooling earlier and, as a result, to attain a lower level of education.

This latter hypothesis is in keeping with traditional differences in the means by which each sex has attained financial security and status in adulthood. Marriage has been, and to some extent still is, the major route to financial well-being and status for women, while occupational achievement is the major route to status attainment for men. Entry into marriage, therefore, can provide a direct means of status attainment for women in much the way that educational attainment does for men. Until recent years when a large percentage of married women have entered the labor force, educational attainment was instrumental to the status attainment of women only in placing them in a better marriage market. It was often in a

woman's best interest to limit her education and marry once she found a desirable mate; conversely, it was usually in a man's best interest to continue his education to advance his status prospects. Because of these differences in the means by which each sex traditionally has attained status, women not only tend to marry earlier than men, but their earlier age at marriage is more likely to cause them to terminate their schooling. Thus, the fact that women take on the adult roles of spouse and parent earlier may be an important factor limiting their educational and subsequent occupational achievement.

Sex Differences in the Determination of Educational Attainment and Age at Marriage

Although considerable attention has been devoted to developing models of the male educational attainment process (Sewell et al., 1970; Kerckhoff, 1974; Alexander and Eckland, 1974; Alexander et al., 1975; Sewell and Hauser, 1975; Wilson and Portes, 1975; Kerckhoff and Campbell, 1977), relatively little attention has been paid to the study of either the female educational attainment process or sex differences in the educational attainment process. Most studies examining the educational attainment process for both sexes have focused on an intermediate stage in that process, specifically the determination of educational expectations (Sewell and Shah, 1968b; Bayer, 1969; Hauser, 1971; Marini, 1974; Hout and Morgan, 1975; Marini and Greenberger, 1978; Rosen and Aneshensel, 1978). Within the last few years more attention has been given to considering sex differences in the determination of educational attainment, but these analyses usually have been carried out as part of larger studies of sex differences in the status attainment process, and most have been based on comparisons of working women and men and/or married women and men (Carter, 1972; Treiman and Terrell, 1975; Alexander and Eckland, 1975; Featherman and Hauser, 1976; McClendon, 1976; Spaeth, 1977). Since working women are known to attain more education than nonworking women (Bowen and Finegan,

1969:117; Treiman and Terrell, 1975:178) and since comparisons of spouses, such as that carried out by Featherman and Hauser (1976), are conditioned by the effects of assortative mating (Blau and Duncan, 1967:354-9), these studies offer limited insight into the extent and nature of sex differences in the educational attainment process.³

From the small amount of research which has been done on full samples of women and men, several important findings have emerged. One is that, at least for the cohort who attended high school in the late 1950s, socioeconomic background had a stronger effect on the educational attainment of females than males; academic aptitude, on the other hand, had a stronger effect on educational attainment for males than females (Sewell and Shah, 1967; 1968a; Alexander and Eckland,

1974). Similar sex differences in the effects of these two variables on educational expectations (Turner, 1966; Sewell and Shah, 1967; 1968a; 1968b; Bayer, 1969; Hauser, 1971; Alexander and Eckland, 1974; Marini, 1974; Hout and Morgan, 1975), perceived parental encouragement (Sewell and Shah, 1968a; 1968b), and enrollment in a college preparatory curriculum (Alexander and Eckland, 1974) also have been documented. These results suggest that there is somewhat greater status ascription in the educational attainment process for females than males and that, conversely, academic ability plays a more important role in the educational attainment process for males than females.

A second important finding, however, appears to be that when additional variables are included in the analysis, sex differences in the effects of socioeconomic background and ability on intervening variables in the educational attainment process generally are not significant. Alexander and Eckland (1974) found no interaction between sex and either socioeconomic background or ability in the determination of parental encouragement, curriculum, the influence of significant others, academic self-image, and educational expectations; however, they did find significant sex interactions in the determination of senior class standing and educational attainment. Their findings suggest that there may be few sex differences in the process of educational attainment up through the high school years and that it is primarily in the conversion of the academic resources acquired during high school into additional years of schooling that the educational attainment process differs for the two sexes.

Additive sex effects have been found on a number of intervening variables in the educational attainment process, however. These effects indicate that the two sexes differ on the values of the variables in question rather than on the relationships between variables. Several additive sex effects disadvantage women in the attainment process. These include the perception of more parental encouragement by males than females and the more frequent enrollment of males than females in a col-

³ A comparison of two analyses by Alexander and Eckland (1974; 1975), one of which was based on all women in the sample and the other of which was based only on women working at two points in time, indicates that a consistent bias emerges when employed women only are analyzed. Because employed women have attained a higher level of education than women who are not in the labor force, an analysis based only on employed women underestimates the direct additive effect of sex on educational attainment. Restricting the analysis to employed women also affects the detection of interactive sex effects by altering somewhat the relationships among variables for women throughout the model. For example, in studying the effects of socioeconomic background and academic aptitude on each variable in the model, Alexander and Eckland present results showing that, when the analysis is based only on employed women, the metric coefficients for the effects of socioeconomic background on educational attainment and on the variables most strongly mediating the effects of socioeconomic background on educational attainment are underestimated; the metric coefficient for the effect of academic aptitude on educational attainment, on the other hand, is slightly overestimated. In contrast, the standardized coefficients for the effects of academic aptitude are generally underestimated, while those for the effects of socioeconomic background tend to be overestimated except when the corresponding metric coefficient is underestimated. These differences lead to the documentation of statistically significant sex interaction effects in five more equations for the sample which is restricted to employed women than for the unrestricted sample, despite a smaller number of cases in the restricted sample. In other equations for which statistically significant sex interactions are documented on both samples, the magnitude of the sex interaction effect is underestimated on the restricted sample.

lege preparatory curriculum (Alexander and Eckland, 1974; Hauser et al., 1976). Several other sex effects offer advantages to women. Among these are the higher grades earned by females than males and their concomitantly higher academic self-image (Alexander and Eckland, 1974; Hauser et al., 1976). There appears to be no sex difference in the perception that friends will attend college (Alexander and Eckland, 1974; Hauser et al., 1976), and evidence is conflicting with respect to the effect of sex on the perception of teacher encouragement and educational expectations. Alexander and Eckland (1974) found no sex effect on either of these variables with the effects of other variables controlled, but Hauser et al. (1976) found that males more often perceived teacher encouragement and had higher educational expectations.

To some extent, then, the lower educational attainments of women can be understood in terms of additive sex effects on intervening variables in the educational attainment process. Even with the effects of these variables controlled, however, sex continues to exert highly significant additive and interactive effects on educational attainment (Alexander and Eckland, 1974; Hauser et al., 1976). These latter effects remain largely unexplored because variables which are likely to be important in the female educational attainment process have not been included in earlier models. The most important of these variables is probably age at first marriage.

Very little research has been done on the determinants of age at first marriage or on the relationship between age at first marriage and educational attainment for either sex. A number of variables, such as socioeconomic background, farm background, number of siblings, intelligence, academic performance, dating behavior, educational plans, and occupational aspirations, have been identified as related to age at first marriage (Burchinal, 1959; Bayer, 1968; 1969; Bartz and Nye, 1970; Elder, 1972; Elder and Rockwell, 1976; Call and Otto, 1977; Voss, 1977), but the causal nature of these relationships has not been studied adequately, and little attention has been given to sex compari-

sons. The most problematic aspect of previous research on age at marriage is the specification of the causal relationship between age at marriage and educational attainment. Although this relationship is well-documented (U.S. Bureau of the Census, 1969; 1973; Ryder and Westoff, 1971:298; Scanzoni, 1975:63-103), most researchers have left the causal direction of the relationship unspecified, recognizing that it may run in either or both directions (Ryder and Westoff, 1971:298-9; Elder, 1972; Elder and Rockwell, 1976). One recent study has examined the unidirectional effect of age at marriage on educational attainment (Call and Otto, 1977), but by assuming only a one-way effect, it precludes analysis of the causal nature of the relationship. If the direction of the relationship is other than the one assumed, it will overestimate the effect of age at marriage on educational attainment.⁴

Only one previous attempt has been made to estimate a simultaneous-equation model in which age at marriage and educational attainment are hypothesized to reciprocally influence each other (Voss, 1977). That analysis produced the anomalous result that age at first marriage had a negative effect on educational attainment. It seems likely that such a finding arose from the fact that, with the exception of one variable—age at graduation from high school—the exogenous variables used to estimate the model via two-stage least squares had little effect on age at first marriage. During the estimation process, age at graduation from high school therefore, in effect, was substituted for age at first marriage. Since this variable was related negatively to educational attainment, the effect of age at first marriage was also found to be negative.⁵

⁴ Since Call and Otto (1977) found no significant effect of age at marriage on educational attainment for males, changing their specification would not alter this finding.

⁵ Voss's (1977) analysis was based on the 1957 sample of Wisconsin high school seniors studied by Sewell and his associates (Sewell and Shah, 1967; 1968a; 1968b; Sewell and Hauser, 1975). Although the measures of educational attainment and age at first marriage were taken only seven years after high school and the analysis therefore excludes those marrying at later ages, characteristics of the sample

The correlations upon which this analysis was based indicate that the relationship between age at marriage and educational attainment is stronger for females than males, but the reciprocal effects between these two variables require the consideration of other variables in order to be estimated adequately for each sex.

DATA AND METHODS

This analysis is based on data from a study of students in ten Illinois high schools in 1957-58 and a fifteen-year follow-up survey of the same individuals in 1973-74. The original study was carried out by Coleman (1961) and reported in his book, *The Adolescent Society*. The purpose of Coleman's analysis was to investigate the nature and consequences of high school status systems. The sample of schools upon which the research was based was selected therefore to provide variation in size, in type of community location, and in adolescent status system. With the exception of a small number of absentees, the sample comprised all students in the ten high schools.⁶ Usable data were obtained from a total of 8,617 students (4,352 females and 4,265 males) via a questionnaire administered in the fall of 1957⁷ and from 64% of these students' parents via a mailed questionnaire. Data were also collected from the students' school records.

The follow-up survey was designed and carried out fifteen years after the original

data collection period by Temme⁸ and Marini. Data were obtained from 6,498 (3,433 females and 3,065 males) of the former students, or 75% of the original adolescent sample. The data were obtained initially by mailed questionnaire and then by telephone interview from those who did not return a questionnaire. Forty-four percent of the original sample responded to the questionnaire and 31% to the shorter telephone interview. The rate of refusal to participate was low, the non-respondents primarily being individuals who could not be located.

Although the original sample was a purposive one rather than one chosen to be representative of a geographically-defined population, analyses based on this sample (Marini, 1974) show striking similarity to those based on other samples of this cohort, including the sample of Wisconsin high school seniors analyzed by Sewell and his associates (Sewell and Shah, 1967; 1968a; 1968b; Sewell and Hauser, 1975) and the EEO sample of sophomores in selected U.S. high schools analyzed by Alexander and Eckland and their associates (Alexander and Eckland, 1974; Alexander et al., 1975). A comparison of data obtained in the follow-up survey with data reported for this cohort by the U.S. Census Bureau also shows that the distribution of the sample on major demographic variables, such as educational attainment, employment status, marital status, fertility, age at first marriage, and age at the birth of each child, is quite similar to the distribution for the U.S. population as a whole. The sample is somewhat more highly educated than the U.S. population, however, a difference which stems in part from the fact that those comprising the sample already had completed various stages of high school at the time of the original survey.

Data from the adolescent survey and school records were used to analyze the extent to which respondents to the follow-up and those whose parents responded to the parents' questionnaire

do not appear to explain Voss's results. A comparison of the correlations upon which his analysis was based with those obtained from the fifteen-year follow-up study of high school students analyzed here indicates that although his correlations were attenuated, the basic pattern of relationships among variables was similar.

⁶ The sample considered here is the sample which formed the basis for the follow-up study. Most of Coleman's analysis was based on students in ten public high schools, nine of which were coed and one of which was all-male. Some comparative data also were presented for students in two private coed high schools (Coleman, 1961:77-80). The follow-up study is based on students in the nine public coed high schools analyzed extensively by Coleman and in one of the two private schools (University High) for which some data were presented.

⁷ The data for the one private school were collected in the fall of 1958.

⁸ In addition to collaborating in the design of the questionnaire, Lloyd V. Temme supervised the location of respondents, the collection of data, and part of the data coding.

were representative of the total original sample. These analyses showed that respondents to the follow-up were highly representative of the total original sample and that those whose parents responded to the parents' questionnaire were only slightly less so. The few differences observed, and even these were small, were in the means and standard deviations of variables predictive of educational attainment, such as grade-point average, enrollment in a college preparatory curriculum, intelligence, participation in extracurricular activities, college plans, and friends' college plans. Because a small subset of variables measured during the follow-up were measured only for respondents to the initial questionnaire and not for respondents to the telephone interview, the extent to which these respondents were representative of the original adolescent sample also was investigated. These initial respondents, who comprise 44% of the original sample, were less representative of the total adolescent sample in their means and standard deviations than was the total follow-up sample. They were higher in educational and occupational achievement and on the variables measured during adolescence which are predictive of these outcomes.

Because data were available for four sets of variables which had been measured for varying numbers of respondents, and because there was appreciable bias in the means and standard deviations for one of the sets of variables which would have rendered an analysis based on only those individuals responding to variables in all four sets also biased, a special procedure was used to obtain maximum likelihood estimates of the means and variance-covariance matrix for the total sample on all variables to be included in the analysis. This procedure is described in detail elsewhere (Marini et al., 1977) and involves the estimation of parameters for variables with missing data via regression on the observed data. The estimation is accomplished in a way that does not underestimate variances, a difficulty usually encountered when estimated values are placed on the regression line and the distribution of points about the line is not taken into consideration. The procedure

employed yields maximum likelihood estimates of the means and variance-covariance matrix for the total original sample under the assumption that relationships among variables are similar for respondents and nonrespondents. This assumption is highly plausible, given the similarity of relationships between respondents and nonrespondents which has been documented on this and other similar data sets (Taubman and Wales, 1974:209-15; Sewell and Hauser, 1975:31-41).⁹

THE MODEL

The basic outline of the model postulated to explain educational attainment and age at first marriage is presented in Figure 1. Educational attainment and age at marriage are hypothesized to have reciprocal effects on one another. The major causal influence is expected to be that of educational attainment on age at marriage, but there also are expected to be individuals for whom early marriage limits further education. Educational attainment and age at marriage are both seen as outcomes of influences earlier in the life cycle. Because these influences are ordered in time and in the direction of their effects on one another, a multistage model is required to accurately characterize the causal process by which educational attainment and age at marriage are determined. Five basic types of variables are considered to influence educational attainment and age at first marriage. These are discussed in turn below.

Background influences. The back-

⁹ The appropriateness of the missing data procedure employed could be verified further by comparing the correlations estimated between number of siblings, one of the variables observed for only 44% of the original sample, and other variables with comparable correlations obtained on data sets where the number of siblings was measured for all respondents. When the correlations estimated between number of siblings and father's education, father's occupation, measured intelligence, and educational expectations were compared with corresponding correlations obtained by Hauser (1971), Kerckhoff (1974), and Sewell et al. (1977), the correlations obtained on this data set were found to be similar to those obtained on other data sets, although there were some differences among the correlations obtained on the data sets used to make the comparison.

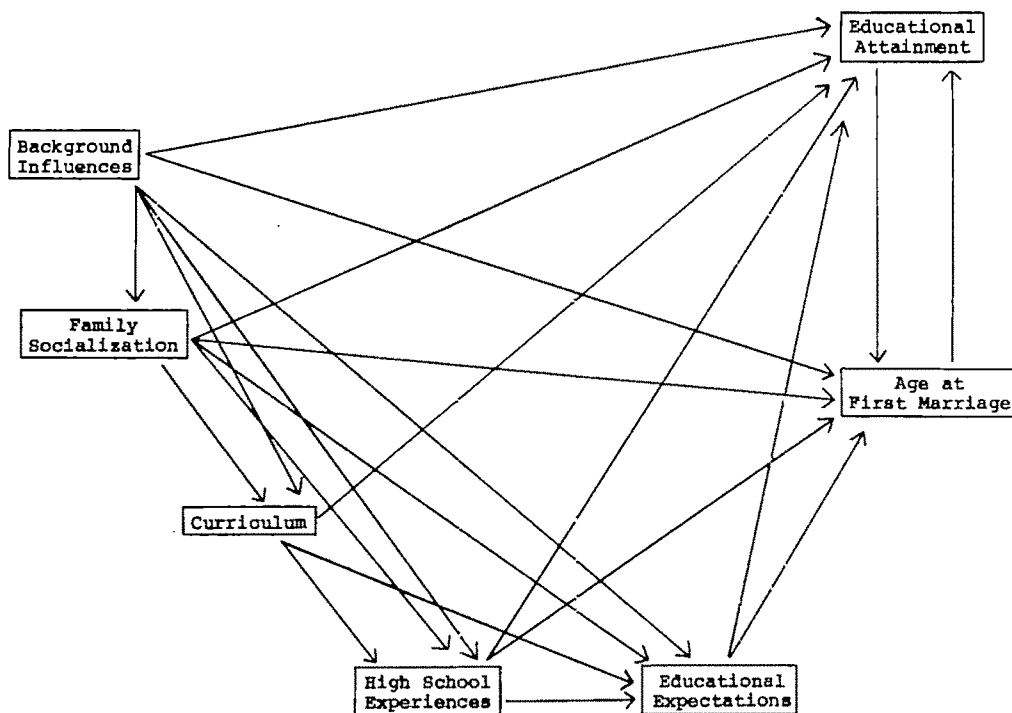


Figure 1. Outline of Model for the Determination of Educational Attainment and Age at First Marriage

ground influences, or exogenous variances, included in the model are those which frequently have been included in models of the educational attainment process. These are measures of socioeconomic background, number of siblings, and intelligence. Mother's education, father's education, father's occupational prestige, and family income were used to measure socioeconomic background. Other background variables also were examined in relation to educational attainment and age at marriage but were not included in the final model because they did not add significantly to the proportion of variance explained by socioeconomic background, number of siblings, and intelligence. These variables were: religion, age, position in the birth order, whether the parents were foreign-born, whether the mother or father died before the respondent was 18 years old, whether the parents were separated or divorced, whether the father was away during World War II, and whether the respondent lived with relatives other than her or his parents during childhood. The fact that these variables were not retained in the model does

not necessarily mean that they did not affect educational attainment or age at marriage. In some cases, especially when an event was rare, the amount of variance in the independent variable was simply too small for the variable to have a significant effect on either dependent variable. These variables were not included in the model because they did not contribute significantly to its explanatory power.

The contextual effect of the high school was also not included in the model since, despite the fact that the ten schools were chosen for their apparent differences, only 9% of the variance in educational attainment was found to lie between schools. With the effects of the student's socioeconomic background and intelligence controlled, differences between schools added only about 1% to the percent of variance in educational attainment explained by the individual-level variables. These findings are remarkably similar to those reported by other investigators (Jencks and Brown, 1975; Hauser et al., 1976).

Family socialization. Only one direct measure of family socialization was in-

cluded in the model. Based on responses obtained from the students' parents, it was an indicator of parental encouragement to go to college. Specifically, the variable measured whether the parents indicated they would like to see their child go to college. Parental encouragement to go to college was hypothesized to be influenced by parental socioeconomic status, number of siblings, and intelligence. Because, for the most part, parental encouragement was regarded as an enduring parental attitude which has an important effect on the respondent's placement in a college preparatory curriculum at entry into high school, no effect of high school grades on parental encouragement was included in the model, as it has been by other investigators (Sewell and Hauser, 1975:91-100; Alexander and Eckland, 1974). Rather, parental encouragement was assumed to affect curricular placement (or the student's college plans at entry into high school), which, in turn, was postulated to affect high school grades. Although it is likely that high school grades do, to some extent, affect parental encouragement to go to college, the dominant influence of parental encouragement is likely to occur earlier, either at or prior to the time of curricular placement, and therefore is not influenced by subsequent high school grades.

A number of other measures of family socialization were also examined in relation to educational attainment and age at first marriage but were not included in the final model because they did not have a significant effect on either of the dependent variables. These variables included measures of the mother's participation in the labor force and attitude toward working, parental rules concerning time spent on homework and watching television, parental participation in community organizations, the relative importance of the mother vs. the father in various aspects of socialization, the closeness of the relationship between the two parents and the relationship between the respondent and each parent, and some general measures of the parents' attitudes toward educational and occupational achievement.

Curriculum. Whether or not the re-

spondent was enrolled in a college preparatory curriculum was the only variable in the third stage of the model. Largely determined at entry into high school, this variable was seen as influenced by both the exogenous background variables and parental encouragement to go to college. In turn, it was seen as influencing experiences during high school, educational expectations, and subsequent educational attainment. Because educational expectations were measured only during the high school years and not at entry into high school, the model examined here does not permit an estimate of the effect of curriculum on educational expectations. Only if educational expectations had been measured prior to exposure to the curriculum would it be possible to overcome the selectivity bias resulting from the effect of college plans on curricular placement. Since a measure of college plans at entry into high school was not available, the estimated effect of curriculum on college plans during high school is actually an estimate of the combined effects of curriculum and college plans at entry into high school on college plans at the time of the adolescent survey.

High school experiences. Measures of both academic and nonacademic experiences during high school were included in the fourth stage of the model. These variables were hypothesized to be influenced by the exogenous background variables, parental encouragement to go to college, and curriculum. The variables included in the model were measures of academic effort (the amount of time spent on homework), high school grade-point average, friends' college plans, participation in extracurricular activities, dating frequency, and adolescent rebellion (the frequency of smoking and drinking). With the exception of a direct effect of academic effort on grades and a direct effect of one measure of dating frequency on another, causal relationships were not postulated among these variables.

Two measures of dating frequency were incorporated in the model. One measured the frequency of dating at the time of the adolescent survey, and the second was a retrospective measure taken during the follow-up to determine the frequency of

dating during the last two years of high school. The frequency of dating at the time of the adolescent survey was assumed to affect the frequency of dating during the last two years of high school since the adolescent survey preceded the last two years of high school for more than three-fourths of the respondents and preceded at least the last year of high school for the remaining respondents. Both measures were included in the model in order to obtain more precise estimates of the effects of dating on educational expectations at the time of the adolescent survey and on age at first marriage.

Dating frequency was expected to have a direct negative effect on the educational expectations of females because it is indicative of the likelihood that a relationship had been established with a potential marriage partner and, therefore, of the availability of the opportunity to marry. Since marriage has been a predominant means of status attainment only for women, frequency of dating in high school was expected to affect negatively the educational expectations of females but not those of males. Irrespective of its effect on educational expectations, however, frequency of dating was hypothesized to have a direct negative effect on age at first marriage for both sexes, both because it suggests that a potential marriage partner has been found and because it affects the likelihood that a premarital pregnancy will occur at an early age and precipitate an early marriage. It was also expected that the frequency of dating and educational expectations might interact in their effect on age at first marriage; specifically, that frequency of dating would affect age at first marriage only for those who did not plan to go to college. This interaction effect was not found to be significant, however, and it therefore was not included in the final model.

Like frequency of dating, rebellion against the norms of the adolescent culture, which was measured by the frequency of smoking and drinking, was expected to affect negatively educational attainment via its negative effect on educational expectations. Those defying adult authority by altering the accepted timetable for engaging in adult behavior were

expected to be less likely to want to continue in the subordinate student role. On the contrary, active participation in extracurricular activities was seen as integrating a student into the school milieu, making the student not only more involved with current school life but more likely to desire and expect to continue schooling after high school.

A number of other variables measuring experiences and attitudes during high school were also examined in relation to educational attainment and age at marriage and were not found to contribute significantly to the explanation of variance in either variable once the above variables were taken into consideration. These variables included age at the respondent's first date, age when the respondent first "went steady," whether the respondent ever went steady or was going steady at the time of the adolescent survey, frequency of church attendance, participation in clubs outside of school, whether the respondent had a job at the time of the adolescent survey, the number of hours worked per week, measures of self-esteem, measures of status in the adolescent peer culture, and preferred subjects in high school. Again, it should not be inferred that all of these variables had no effect on the two dependent variables since whether they explained a significant amount of variance was in part a function of their variances.

Educational expectations. With the exception of academic effort, which was assumed to affect educational expectations only via its effect on academic performance, the high school experiences included in the model were postulated to affect directly educational expectations. Only one of the two measures of dating frequency, that taken at the time of the adolescent survey, was assumed to affect educational expectations, also measured at the time of the adolescent survey. In addition, prior influences represented by the background variables, parental encouragement, and curriculum were assumed to have direct and/or indirect effects on educational expectations.

Because the relationship between friends' educational expectations and the respondent's own educational expecta-

tions is represented in the model as a unidirectional effect of friends' expectations on those of the respondent, this effect may be slightly overestimated. Several studies have suggested that a more appropriate specification of the relationship between these two variables is one in which they reciprocally influence each other (Duncan et al., 1968; Hout and Morgan, 1975). These studies, however, did not control for the effect of curriculum when estimating the effect of friends' educational expectations on the respondent's educational expectations and vice versa. By looking at the allocative role played by curriculum in determining the educational expectations of friends, and by controlling for curriculum in estimating the effect of friends' educational expectations on the respondent's educational expectations, it is likely that the effect of friends' educational expectations on the respondent's educational expectations is only slightly overestimated. Since the major emphasis in this analysis is on estimating the reciprocal effects between age at first marriage and educational attainment and since accurate estimates of the effects of friends' educational expectations and the respondent's educational expectations on educational attainment can be obtained under the specification of a unidirectional relationship between these two variables, this specification was adopted.

Age at first marriage and educational attainment. In the final stage of the model reciprocal effects were postulated between age at first marriage and educational attainment, and each of these variables was also assumed to be affected by other variables in the model. In order to identify the model, curriculum was assumed initially to have no direct effect on age at first marriage. It turned out, however, that most of the variables predictive of educational attainment, including socioeconomic background, parental encouragement, curriculum, grade-point average, friends' educational expectations and the respondent's educational expectations, had no direct effect on age at first marriage when the effect of educational attainment was taken into consideration. Similarly, the measure of dating frequency during the last two years of high school,

which proved to be a major predictor of age at first marriage, had no direct effect on educational attainment. Thus, in estimating the reciprocal effects of educational attainment and age at marriage and the effects of other variables on these two outcomes in the final model, identification of the model proved to be no problem.

Measurement and estimation. Details concerning the measurement of each of the variables included in the model are presented in Table 1. Estimation of the effects on all endogenous variables in the model other than age at first marriage and educational attainment was accomplished by ordinary least-squares regression. Estimation of the effects on age at first marriage and educational attainment was accomplished by two-stage least squares (Goldberger, 1964; Johnston, 1972).

FINDINGS

Estimates of effects in the final model for the determination of educational attainment and age at first marriage are presented for each sex in Tables 2A, 2B, and 3. Effects which were examined initially and found to be less than twice their estimated standard error¹⁰ for both sexes were deleted from the final model. Table 2A contains the metric, or unstandardized, coefficients for the direct effects retained in the model. Table 2B contains the corresponding standardized coefficients. Table 3 contains both the metric and standardized coefficients for the total (direct plus indirect) effects of variables in the model on educational expectations, educational attainment, and age at first marriage. In Tables 2B and 3, a sheaf coefficient (Heise, 1972) is presented also, which summarizes the effects of the four indicators of socioeconomic background (mother's education, father's education, father's occupational prestige, and family income) on each endogenous variable. Al-

¹⁰ Because missing data were estimated to obtain maximum likelihood estimates of the mean vector and variance-covariance matrix, the standard errors obtained are technically estimates of the true standard errors. The N used to obtain the estimates was the number of cases for which data were available on all four sets of variables used in the analysis (1,594 females and 1,083 males).

Table 1. Measurement of Variables in the Model

Variable Label	Variable Name	Units of Measurement	Source*	Description	Mean		Standard Deviation	
					Females	Males	Females	Males
FAED	Father's education	Years of schooling completed	AQ		11.702	11.659	3.528	3.454
MOED	Mother's education	Years of schooling completed	AQ		11.508	11.694	2.947	2.847
FAOCC	Father's occupational prestige	NORC occupational prestige scores	F	The NORC scores were reconciled to the 1970 Census occupation and industry codes for detailed titles (Temme, 1973)	44.000	43.884	14.819	14.255
FAMINC	Family income	A scale ranging from 1 to 7	PQ		4.092	4.067	1.528	1.514
SIBS	Number of siblings	Persons	FQ		2.267	2.282	1.929	1.968
IQ	Measured intelligence	Standard IQ scale	REC		106.625	105.761	12.910	14.235
PENC	Parental encouragement to go to college	Dummy variable	PQ	Parents indicated whether or not they would like to see their son or daughter go to college	.714	.824	.435	.360
CURRIC	Enrollment in a college preparatory curriculum	Dummy variable	REC		.411	.458	.492	.498
FEDEXP	Friends' educational expectations	Proportion of friends listed by the respondent who indicated they planned to go to college	AQ					
EFFORT	Academic effort	Hours per day spent studying	AQ		.512	.568	.369	.381
GPA	Grade point average	Usual grading scale ranging from 0 to 4	REC		1.589	1.145	.814	.807
PARTIC	Participation in extracurricular activities	Dummy variable	AQ	Measures active participation based on three questionnaire items	2.324	2.013	.773	.772
					.413	.303	.492	.460

Table 1. Continued

Variable Label	Variable Name	Units of Measurement	Source *	Description	Mean		Standard Deviation	
					Females	Males	Females	Males
REBEL	Adolescent rebellion	A scale ranging from 1 to 3	AQ	Measures frequency of smoking and drinking				
DATE 1	Dating frequency at the time of the adolescent survey	Dates per week	AQ		.570	1.016	1.032	1.425
DATE 2	Dating frequency during the last two years of high school	A scale ranging from 1 to 5	FQ		4.030	2.538	4.802	3.753
EDEXP	Educational expectations	Dummy variable	AQ	Measures whether respondent planned to go to college	3.102	3.096	1.213	1.198
AGEMAR	Age at first marriage	Years	F		.488	.540	.500	.498
EDUC	Educational attainment	Years of schooling completed	F	Includes postsecondary schooling not acquired in a college or university	20.814	22.789	2.816	2.952
					13.625	14.495	2.294	2.864

* AQ = Adolescent questionnaire.

F = Follow-up questionnaire and telephone interview.

PQ = Parents' questionnaire.

FQ = Follow-up questionnaire.

REC = School records.

Table 2A. Metric Coefficients Measuring Direct Effects in a Model of Educational Attainment and Age at First Marriage, By Sex (Standard Errors in Parentheses).

Independent Variable	Dependent Variable															
	PENC	CURRIC	FEDEXP	EFFORT	GPA	PARTIC	REERI	DATE 1	DATE 2	EDEXP	AGEMAR	EDUC				
	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M
FAED	.020 (.004)	.016 (.004)	.015 (.004)	.012 (.006)	.012 (.003)	.012 (.004)	.026 (.010)			.014 (.003)	.014 (.004)	.044 (.016)	.039 (.023)			
MOED	-.012 (.004)	-.007 (.005)	.014 (.003)	.003 (.006)	.008 (.003)	.007 (.004)	.012 (.011)			.013 (.004)	.008 (.005)	.041 (.017)	.048 (.025)			
FAOCC																
FAMINC	.044 (.009)	.039 (.009)	.026 (.012)	.023 (.007)	.062 (.009)	.053 (.018)	.089 (.021)									
SIRS	.001 (.005)	-.011 (.005)	-.021 (.007)	-.012 (.004)	-.008 (.005)	-.030 (.010)	-.024 (.012)									
IQ	.008 (.001)	.005 (.001)	.007 (.001)	.008 (.001)	.003 (.001)	.003 (.001)	-.005 (.002)									
PENC																
CURRIC																
FEDEXP																
EFFORT																
GPA																
PARTIC																
REBEL																
DATE 1																
DATE 2																
EDEXP																
AGEMAR																
EDUC																

* R² cannot be obtained when two stage least squares is used (Hout, 1977). * F=females; M=males.

Table 2B. Standardized Coefficients Measuring Direct Effects in a Model of Educational Attainment and Age at First Marriage, By Sex.

Independent Variable	Dependent Variable																								
	PENC		CURRIC		FEDEXP		EFFORT		GPA		PARTIC		REBEL		DATE 1		DATE 2		EDEXP		AGEMAR		EDUC		
	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	
PAED	.161	.149	.105	.081	.090	.112	.051	.110						-.049	-.122			.101	.095			.067	.047		
MOED	.078	-.056	.087	.029	.066	.052	.066	.041										.078	.046			.052	.048		
FAOCC							.057	.026																	
FAMINC	.154	.165	.080	.086	.255	.211	.100	.167																	
(SES)	(.347)	(.255)	(.238)	(.174)	(.407)	(.350)	(.190)	(.283)						(-.049)	(-.122)			(.162)	(.129)			(.171)	(.157)		
SIBS	.003	-.038	-.084	-.111	-.060	-.041	-.070	-.059						-.057	-.035			-.050	-.065						
IQ	.227	.208	.177	.216	.110	.123	-.064	-.043	.497	.507				-.068	-.096			.164	.196			.120	.060		
PENC			.276	.253	.084	.100	.092	.043						-.113	-.031			.377	.384			.109	.056		
CURRIC					.198	.176	.159	.201	.214	.234	.219	.132	-.128	-.106	-.128	-.113	-.031		.167	.157			.094	.112	
FEDEXP									.117	.136															
EFFORT																									
GPA																		.039	.109			.207	.395		
PARTIC																		.058	.037						
REBEL																		-.009	-.053						
DATE 1																									
DATE 2																.438	.329	-.069	.002						
EDEXP																									
AGEMAR																									
EDUC																									

^aF = females; M = males.

Table 3. Metric and Standardized Coefficients Measuring Total Effects on Educational Expectations, Age at First Marriage and Educational Attainment, by Sex

Independent Variable	Dependent Variable											
	METRIC COEFFICIENTS						STANDARDIZED COEFFICIENTS					
	EDEXP F ^a	EDEXP M	AGEMAR F	AGEMAR M	EDUC F	EDUC M	EDEXP F	EDEXP M	AGEMAR F	AGEMAR M	EDUC F	EDUC M
FAED	.030	.030	.062	.041	.090	.115	.211	.206	.078	.048	.138	.138
MOED	.026	.009	.056	.023	.080	.064	.151	.050	.058	.022	.103	.064
FAOCC	.000	.000	.001	.000	.001	.001	.010	.004	.004	.001	.007	.004
FAMINC	.041	.042	.153	.120	.220	.338	.124	.128	.083	.062	.147	.178
(SES)							(.440)	(.357)	(.203)	(.118)	(.347)	(.343)
SIBS	-.025	-.035	-.030	-.032	-.042	-.090	-.096	-.140	-.020	-.021	-.036	-.062
IQ	.007	.008	.015	.013	.048	.076	.188	.239	.069	.061	.269	.375
PENC	.345	.452	.463	.535	.657	1.502	.300	.327	.072	.065	.125	.189
CURRIC	.448	.452	.813	.550	1.124	1.531	.441	.452	.142	.093	.241	.266
FEDEXP	.226	.205	.499	.379	.730	1.070	.167	.157	.065	.049	.117	.142
EFFORT	.003	.009	.048	.071	.070	.201	.005	.015	.014	.020	.025	.057
GPA	.025	.070	.431	.547	.630	1.543	.039	.109	.118	.143	.212	.416
PARTIC	.059	.040	.026	.016	.038	.044	.058	.037	.004	.002	.008	.007
REBEL	-.004	-.018	-.002	-.007	-.003	-.021	-.009	-.053	-.001	-.004	-.001	-.010
DATE 1	-.007	.000	-.054	-.030	-.020	-.003	.069	.002	-.092	-.039	-.042	-.004
DATE 2			-.464	-.291	-.140	-.029			-.199	-.118	-.074	-.012
EDEXP			.439	.394	.645	1.113			.078	.066	.140	.193
AGEMAR					.302	.102					.371	.106
EDUC			.684	.354					.557	.344		

^a F = females; M = males.

though the sheaf coefficient does not measure the effect of a linear function of the four indicators on the endogenous variables, it does provide a convenient way of representing the influence of the set of variables on a given endogenous variable. The relatively high correlations among the four indicators of socioeconomic background render estimates of the effects of each variable considerably lower than they would be if the variable were entered into the equation without the other three (Gordon, 1968). Attempts to determine the influence of socioeconomic background relative to other variables therefore require some way of summarizing the combined effects of the four indicators. Finally, Table 4 presents the cumulative percentages of variance in age at first marriage and educational attainment explained at various stages of the model. Beginning with the percentage of variance explained by the reduced-form equation, the percentage of variance is presented for each chronological stage of the model and increases as additional variables are added. In discussing the results presented in Tables 2A to 4, attention is focused primarily on sex differences in the effects of variables on educational attainment and age at first marriage and on the relationship between these two variables.

Sex Differences in the Determination of Educational Attainment

Estimates of the total effects of the variables in the model on educational at-

tainment, which are presented in Table 3, indicate three fundamental sex differences in the determination of educational attainment: (1) a greater role modeling effect of mother's education on females than males; (2) greater returns on academic resources and accomplishments in high school for males than females; and (3) an effect of dating frequency and age at first marriage on females but not males. Each of these three sex differences will be examined in turn. The effect of mother's education will be considered first, along with the effects of the other indicators of socioeconomic background.

Socioeconomic background. Estimates of the total effects of the socioeconomic background variables on educational attainment indicate that socioeconomic background has a similar overall effect on the two sexes. Contrary to the findings of previous research (Sewell and Shah, 1967; 1968b; Alexander and Eckland, 1974), the overall effect of socioeconomic background is not stronger for females than males. Of the four indicators, mother's education has a slightly stronger effect on females than males, but the effects of father's education and family income are slightly stronger for males than females. Father's occupational prestige has little effect on educational attainment for either sex. About half of the total effect of socioeconomic background is direct, and the other half is exerted indirectly through other variables in the model. There is no sex difference in the direct effect, and the indirect effects are sometimes greater for

Table 4. Cumulative Percentages of Variance in Age at First Marriage and Educational Attainment Explained at Various Stages of the Model, by Sex

Independent Variables	Percentage of Variance Explained			
	AGEMAR		EDUC	
	Females	Males	Females	Males
Exogenous variables (socioeconomic background, number of siblings, intelligence)	7.8%	4.2%	42.9%	41.6%
Parental encouragement	8.7	5.1	45.1	44.4
Curriculum	10.5	5.6	51.1	49.8
High school experiences	26.0	12.2	60.6	62.4
Educational expectations	26.5	12.4	62.0	63.9
Age at first marriage			a	a
Educational attainment	a	a		

* R^2 cannot be obtained when two-stage least squares is used (Hout, 1977).

females and sometimes greater for males. As shown in Table 2, socioeconomic background has stronger direct effects on parental encouragement, curriculum, friends' educational expectations and the respondent's educational expectations for females than males, but stronger direct effects on academic effort and adolescent rebellion for males than females.

The effects of the socioeconomic background variables on educational attainment generally appear to be of two types: effects attributable to the general level of family social standing, on the one hand, and effects attributable to sex-linked role modeling and socialization, on the other hand. The effects of father's education and family income on variables in the model are usually about equal for the two sexes or greater for males than females; the effect of mother's education, on the other hand, is consistently greater for females than males. Since the general position of the family within the social hierarchy is tied more strongly to the father's education (Rossi et al., 1974) and family income, it seems reasonable to conclude that the finding of no sex difference in the overall effect of socioeconomic background on educational attainment is generated by two offsetting sex differences: a somewhat greater effect of family social standing, or the general level of resources, on males than females and a greater sex-linked modeling effect of mother's education on females than males. While there also may be a somewhat greater modeling effect of father's education on males than females, the sex difference in this effect is, in general, smaller than the sex difference in the effect of mother's education, either because the sex-linked modeling effect is confounded by the relationship between father's education and family social standing or because there is greater cross-sex modeling among females than males.

Academic resources and accomplishments in high school. Estimates of the total effects of most of the other variables in the model which significantly affect educational attainment for at least one sex are greater for males than females. The variables having greater total effects for males than females include number

of siblings, intelligence, parental encouragement, curriculum, friends' educational expectations, academic effort, grades, and the respondent's educational expectations. The sex differences in the effects of number of siblings, curriculum, and friends' educational expectations are too small to be of any consequence, but the overall pattern of greater effects for the set of variables on males than females indicates that males generally get a higher return on their academic resources and accomplishments in high school than females. Resources such as intelligence and parental encouragement, the investment of effort in studying, grade attainment, and the intention to go to college all are linked more strongly to educational attainment for males than females.

The total effect of measured intelligence on educational attainment is about equal to the total effect of socioeconomic background for males but smaller than the effect of socioeconomic background for females. Most of the effect is indirect, but there is also a significant direct effect for both sexes. The direct effect is about twice as strong for females (.021) as males (.012), and the indirect effect is almost three times as strong for males (.064) as females (.027). The indirect effect is exerted primarily through parental encouragement, curriculum, grades, and friends' educational expectations for both sexes, and through frequency of dating for females.

The total, direct, and indirect effects of parental encouragement on educational attainment are also greater for males than females. The total effect of parental encouragement is more than twice as great for males as females, resulting in an additional one and a half years of schooling for males and only an additional .6 years of schooling for females. It is difficult to interpret this effect because the measure of parental encouragement is primarily a measure of the parents' attitude toward the child's going to college. It seems likely that whether or not parents want their child to go to college has a greater effect on the educational attainments of males than females at least in part because the parents' desire is accompanied more often by explicit encourage-

ment and support when the child is a male than a female. This hypothesis is consistent with the fact that a higher proportion of the parents of male than female students even expressed the desire that their child go to college. The indirect effect of parental encouragement on educational attainment is exerted through curricular placement, friends' educational expectations, academic effort, and educational expectations.

Academic effort, or the amount of time spent on homework in high school, was not assumed to have a direct effect on educational attainment but to affect educational attainment indirectly through its effect on high school grades. Although the effect of academic effort on grades is about the same for the two sexes, grades have a stronger effect on educational attainment for males. The indirect effect of academic effort, mediated by grades, is therefore stronger for males than females. For every additional hour per day spent studying in high school, males attained an additional .2 years of schooling. In contrast, there was no significant effect of studying on the educational attainment of females. Part of this sex difference seems to be explained by the fact that females generally spend more time studying than males and their studying is less influenced by variables predictive of future educational plans. Both socioeconomic background and enrollment in a college preparatory curriculum had stronger effects on the studying of males than females. Although the studying of females was influenced somewhat more strongly by parental encouragement and intelligence than that of males, in general, variables in the model were more predictive of studying for males than females.

Both the total and direct effects of high school grades on educational attainment were more than twice as strong for males as females. A one point increase in grade-point average resulted in an average of one and a half additional years of schooling for males but only .6 additional years of schooling for females. Almost all of this effect was direct for both sexes, since relatively little of it was mediated through educational expectations. The effect of grades on educational expectations

was greater for males than females, however; for females, it was barely significant. Again, these sex differences in the effect of grades to some extent reflect differences in the academic orientations of females and males. During high school the academic performance of females is generally higher than that of males, in part because they spend a greater amount of time studying. Not only was the mean grade-point average higher for females than males, but there was also less within-person variance in grades for females than males (Coleman, 1961). The irony, of course, is that despite their superior performance in high school, women are less likely to go to college and attain successively higher levels of education.

Plans to attend college have a direct effect on educational attainment for both sexes, although the size of the effect is almost twice as great for males as for females. For males, planning to go to college results in an additional 1.1 years of schooling, while for females it results in an additional .6 years of schooling. This difference indicates that plans to go to college are less predictive of total educational attainment for females than males, but whether this difference is due to the fact that females are less likely to realize their goal of going to college or to the fact that females attain less education beyond college cannot be determined from this analysis. The analysis does show, however, that the intention to go to college ultimately leads to a higher level of educational attainment for males than females.

Neither participation in extracurricular activities nor adolescent rebellion has an effect on educational attainment. Being an active participant in school-related activities has a slight positive effect on the educational expectations of females, but no effect on subsequent educational attainment. The frequency of smoking and drinking, which is interpreted here as a form of rebellion against the norms of adolescent behavior, has a small negative effect on the educational expectations of males, but, again, no effect on educational attainment.

Dating frequency and age at first marriage. The only two variables, aside from

mother's education, which have greater total effects on educational attainment for females than males are dating frequency in high school and age at first marriage. Neither of these variables has a significant effect on males. This sex difference reflects the greater importance of marriage to the financial security and status attainment of women. Since marriage has been an important means of status attainment for women but not for men, dating frequency, which affects the timing of marriage, and the timing of marriage itself affect the educational attainments of women but not men.

Frequency of dating in high school has a negative effect on both the educational expectations and attainments of females. All of the effect of dating frequency on educational attainment is indirect and mediated through educational expectations and age at first marriage. Although age at first marriage rarely has been considered in earlier models of the educational attainment process, its effect on educational attainment is greater for females than that of any other variable in the model. A difference of one year in age at marriage makes a difference of .3 years in schooling for females.

Explanation of the variance in educational attainment. As indicated in Table 4, all of the exogenous variables in the model, including the four measures of socioeconomic background, number of siblings, and intelligence, together explain 42.9% of the variance in educational attainment for females and 41.6% of the variance in educational attainment for males. Although the effect of socioeconomic background is about equal for the two sexes and the effects of number of siblings and intelligence are somewhat greater for males than females, these variables explain slightly less of the variance in educational attainment for males because the amount of variance in the dependent variable is greater for males than females. That more than 40% of the variance in educational attainment is explained by these exogenous variables for both sexes indicates that a sizeable proportion of the variance in educational attainment is explained by factors which are more or less fixed early in life and are

generally beyond the control of the individual.

Parental encouragement adds about 2% to the percent of variance in educational attainment explained by the exogenous variables for females and about 3% to the percent of variance explained for males. Curriculum, which, as mentioned above, includes both the effect of curriculum and the effect of college plans at entry into high school, in turn adds another 6% for females and 5% for males to the percent of variance in educational attainment explained by the exogenous variables and parental encouragement. Together, the exogenous variables, parental encouragement, and curriculum explain 51.1% of the variance in educational attainment for females and 49.8% of the variance in educational attainment for males. Other experiences during high school add about 10% for females and about 13% for males to the percent of variance in educational attainment explained by the exogenous variables, parental encouragement, and curriculum. Educational expectations in turn add another 1.5% for each sex to the percent of variance in educational attainment already explained by other variables in the model. Variables measuring influences and expectations during or prior to the high school years therefore account for a total of 62.0% of the variance in educational attainment for females and 63.9% of the variance in educational attainment for males.

Because the effect of age at first marriage on educational attainment was estimated using two-stage least squares, it is not possible to separate the variance in educational attainment into explained and unexplained parts (Hout, 1977) and thereby determine the increment in the percent of variance explained by adding age at marriage to the model. The importance of age at first marriage in the educational attainment process for females, however, is underscored by the fact that the effect of age at first marriage is greater than the total effect of any other variable in the model, including the combined effect of the indicators of socioeconomic background. The striking difference between this finding and the finding that age at marriage has no significant effect on

educational attainment for males suggests that causal models developed to explain the behavior of males should not be assumed to represent accurately causal processes for females or to provide an appropriate means of comparing the effects of variables on outcomes for the two sexes. Since the estimated effect of a variable can be biased by the omission of other relevant variables, the use of male models to examine and compare causal processes influencing the behavior of the two sexes can lead to incorrect conclusions.

The analysis presented here focuses on sex differences in relationships among variables in the educational attainment process. Sex differences in educational attainment, however, not only arise from sex differences in the effect of one variable or another, but also from sex differences in the mean values of the variables influencing educational attainment. Analyses on this data set reported elsewhere, which examine the relative importance of these two types of differences in accounting for the sex difference in educational attainment, indicate that sex differences in relationships among variables explain about twice as much of the sex difference in educational attainment as sex differences in the mean values of the variables influencing educational attainment (Marini, 1977a). The implication of this finding is that the lower educational attainment of women than men is due primarily to the effect of sex-role differentiation in adulthood on the conversion of high school resources into educational attainments. Although some of the difference is due to additive sex effects on the parents' desire that the child attend college, enrollment in a college preparatory curriculum, friends' and own college plans, and age at first marriage, more of the difference is due to the differential importance of educational attainment to fulfillment of traditional sex roles and the effect of that difference on the pursuit of higher education by women and men with equal levels of resources.

Sex differences of the type documented here are to be expected in a social system that places a higher value on educational and occupational achievement for males

than females. Since occupational achievement is more important to fulfillment of the male than female role, males who possess high levels of resources which can be applied to the attainment of higher education and, consequently, to the attainment of higher occupational prestige can be expected to do so. Females with similarly high levels of resources, however, are less likely to strive for higher education because their anticipated rewards for doing so are lower. The relationship between the level of academic resources a student has available in high school and educational attainment is therefore weaker for females than males. Conversely, the relationships between educational attainment and variables such as dating frequency and age at first marriage, which affect movement into family roles that have been more important to fulfillment of the traditional female role, are stronger for females than males.

Sex Differences in the Determination of Age at First Marriage

Few variables in the model exert direct effects on age at first marriage. Only educational attainment, dating frequency in high school, and intelligence directly affect age at first marriage for either sex. The other variables in the model that affect age at first marriage do so indirectly, primarily through their effects on educational attainment and, secondarily, through their effects on the frequency of dating in high school. Socioeconomic background, intelligence, parental encouragement, curriculum, friends' educational expectations, grade-point average, and own educational expectations all have positive indirect effects on age at first marriage, which are mediated primarily through educational attainment.

Educational attainment is the most important influence on the timing of marriage for both sexes, but it has a stronger effect on females than males. Each additional year of education results in an increase of .68 years in age at marriage for females but only .35 years for males. This difference arises primarily from the tendency of women to marry at younger ages than men. Since women marry at younger

ages, their education and the timing of their marriages tend to directly impinge upon one another to a greater extent. Women who do not pursue higher education tend to marry shortly after high school, while men tend to marry later regardless of their educational behavior.

Frequency of dating exerts a negative effect on age at first marriage, both directly and indirectly via educational expectations and educational attainment. These effects are again greater for females than males. Thus, those who date more often in high school tend to marry earlier, but the likelihood that frequent dating will lead to early marriage is greater for females than males. Finally, intelligence exerts a direct negative effect on age at first marriage for both sexes, an effect which is also slightly stronger for females than males. When the effect of educational attainment is controlled, those higher in intelligence show a slight tendency to marry earlier. This effect may in part reflect the tendency of those with higher intelligence to graduate from high school somewhat earlier.

In general, the stronger effects of educational attainment, dating frequency, and intelligence on age at first marriage for females than males result from the tendency of women to marry at younger ages than men. Each of these variables reflects the opportunity to marry—specifically, the availability of time free from the requirements of schooling (educational attainment and intelligence) and the availability of a potential mate (dating frequency in high school). Since there is greater social pressure on females than males to marry as a means of attaining financial security and status in adulthood, the age at which women marry is more strongly related to the opportunity to do so than is true for males.

Because the effect of educational attainment on age at first marriage is considerably stronger for females than males, other variables in the model which affect age at marriage have either stronger total effects for females than males or similar total effects for the two sexes. In the latter case, the stronger effect of a variable on educational attainment for males than

females is counterbalanced by the stronger effect of educational attainment on age at first marriage for females than males.

Explanation of the variance in age at first marriage. The exogenous variables in the model explain 7.8% of the variance in age at first marriage for females but only 4.2% of the variance in age at first marriage for males. Parental encouragement adds about 1% for both sexes to the percent of explained variance, and curriculum adds another 2% for females but only .5% for males. Together, the exogenous variables, parental encouragement to go to college, and curriculum explain almost twice as much of the variance in age at first marriage for females as males: 10.5% of the variance is explained for females but only 5.6% of the variance for males.

High school influences and educational expectations also explain more of the variance in age at first marriage for females than males. For females, they add 16% to the percent of variance in age at first marriage explained by the exogenous variables, parental encouragement to go to college, and curriculum; for males, on the other hand, they add only about 7% to the percent of variance explained by these other variables. Thus, variables measuring influences or expectations during or prior to the high school years explain a total of 26.5% of the variance in age at first marriage for females but only 12.4% of the variance in age at first marriage for males. This sex difference in the amount of variance explained is due to the effect of variables in the model on educational attainment, which in turn has a greater effect on age at first marriage for females than males. Again, it is not possible to determine the increment in the percent of variance explained by adding educational attainment to the model since two-stage least squares was used to estimate the effect of educational attainment on age at first marriage. Educational attainment, however, has an effect which is more than twice as large as the total effect of any other variable in the model for both sexes and is therefore likely to add considerably to the amount of variance explained.

CONCLUSION

The traditional differentiation of adult roles by sex has important implications for the process of transition to adulthood. Although educational attainment is the most important variable mediating the transition to adulthood for both sexes and, as such, strongly affects the timing of entry into adult work and family roles, the relationship between educational attainment and the timing of entry into family roles differs for the two sexes. Because marriage has been a major route by which women attain financial security and status in adulthood, early marriage has an inhibitory effect on women's educational attainment. For men, there is no such effect since occupational attainment is the major route by which men attain status in adulthood, and educational attainment is an important prerequisite for occupational attainment.

There is also an effect of educational attainment on age at first marriage for both sexes. This effect is stronger for women than the reciprocal effect of age at first marriage on educational attainment. Moreover, the effect for women is stronger than the effect for men because women tend to marry earlier than men. Thus, women who do not pursue higher education tend to marry shortly after high school, whereas men tend to marry later regardless of their educational behavior.

Just as the relationship between educational attainment and age at first marriage differs for the two sexes, the effects of other variables on these outcomes also differ. Variables such as family income, intelligence, parental encouragement to go to college, academic effort, grades, and the respondent's educational expectations, which represent resources in high school that can be applied to the attainment of higher education, have stronger effects on the educational attainment of males than females. Only mother's education, dating frequency in high school, and age at first marriage have stronger effects on the educational attainment of females than males. Mother's education exerts a greater sex-linked modeling effect on females, and dating frequency in high

school has an inhibitory effect on educational attainment, which is mediated through its effect on college plans and age at first marriage. Both dating frequency in high school and age at first marriage affect the educational attainment of females only, effects which arise from the importance of marriage as a means of status attainment for women. Dating frequency in high school and educational attainment, which are the major determinants of age at first marriage, also have stronger effects on age at first marriage for females than males. These variables reflect the availability of the opportunity to marry since dating frequency is indicative of the likelihood that a potential marriage partner has been found, and educational attainment is indicative of the age at which an individual is free from the requirements of schooling and therefore able to marry.

Thus, variables that measure resources for educational attainment have stronger effects on the educational attainment of males than females, and variables that are indicative of the availability of the opportunity to marry have stronger effects on the age at first marriage for females than males. Just as early marriage has an inhibitory effect on educational attainment only for females, sex differences in the effects of these other variables on educational attainment and age at first marriage reflect the differential importance of educational attainment and marriage to the status attainment processes of the two sexes. Because educational attainment plays a more important role in the status attainment process for men, males are more likely than females to convert their educational resources in high school into subsequent educational attainment. Similarly, because marriage plays a more important role in the status attainment process for women, females are more likely than males to capitalize on the opportunity to marry and to allow early marriage to interfere with their educational attainment. The behavior of both sexes must be viewed as instrumentally realistic given the means of status attainment traditionally open to them. It is the divergent paths of the two sexes upon leaving high school that most strongly differentiate the

processes by which such variables as educational attainment and age at first marriage are determined, a divergence that is not well-predicted by variables measuring behavior earlier in the life cycle but is in accord with the pattern of traditional sex-role differentiation. This pattern, of course, is now undergoing considerable change, and it is to be expected that, as the labor force participation of married women increases and greater emphasis is placed on occupational attainment as the principal means of status attainment for both sexes, the process of transition to adulthood will become more similar for the two sexes. Later cohorts of women are likely to show an increased tendency to convert their educational resources in high school into subsequent educational attainment and a decreased tendency to let marriage interfere with educational attainment.

REFERENCES

- Alexander, K. L. and B. K. Eckland
 1974 "Differences in the educational attainment process." *American Sociological Review* 39:668-82.
 1975 "School experience and status attainment." Pp. 171-210 in S. E. Dragastin and G. H. Elder, Jr. (eds.), *Adolescence in the Life Cycle*. New York: Halsted.
- Alexander, K. L., B. K. Eckland and L. J. Griffin
 1975 "The Wisconsin model of socioeconomic achievement: a replication." *American Journal of Sociology* 81:324-42.
- Bartz, K. W. and F. I. Nye
 1970 "Early marriage: a propositional formulation." *Journal of Marriage and the Family* 32:258-68.
- Bayer, A. E.
 1968 "Early dating and early marriage." *Journal of Marriage and the Family* 30:628-32.
 1969 "Marriage plans and educational aspirations." *American Journal of Sociology* 75:239-44.
- Blau, P. M. and O. D. Duncan
 1967 *The American Occupational Structure*. New York: Wiley.
- Bowen, W. G. and T. A. Finegan
 1969 *The Economics of Labor Force Participation*. Princeton: Princeton University Press.
- Burchinal, L. G.
 1959 "Adolescent role deprivation and high school age marriage." *Marriage and Family Living* 21:378-84.
- Call, V. R. A. and L. B. Otto
 1977 "Age at marriage as a mobility contingency: estimates for the Nye-Berardo model." *Journal of Marriage and the Family* 39:67-79.
- Carter, N. D.
 1972 "The effects of sex and marital status on a social-psychological model of occupational status attainment." Master's thesis, Department of Sociology, University of Wisconsin.
- Coleman, J. S.
 1961 *The Adolescent Society*. New York: Free Press.
- Coleman, J. S., R. H. Bremner, B. R. Clark, J. B. Davis, D. H. Eichorn, Z. Griliches, J. F. Kett, N. B. Ryder, Z. B. Doering and J. M. Mays
 1974 *Youth: Transition to Adulthood*. Chicago: University of Chicago Press.
- Duncan, O. D., A. O. Haller and A. Portes
 1968 "Peer influences on aspirations: a reinterpretation." *American Journal of Sociology* 74:119-37.
- Duncan, O. D., D. L. Featherman and B. Duncan
 1972 *Socioeconomic Background and Achievement*. New York: Seminar Press.
- Elder, G. H., Jr.
 1972 "Role orientations, marital age, and life patterns in adulthood." *Merrill-Palmer Quarterly* 18:3-24.
- Elder, G. H., Jr. and R. C. Rockwell
 1976 "Marital timing in women's life patterns." *Journal of Family History* 1:34-53.
- Featherman, D. L. and T. M. Carter
 1976 "Discontinuities in schooling and the socioeconomic life cycle." Pp. 133-60 in W. H. Sewell, R. M. Hauser and D. L. Featherman (eds.), *Schooling and Achievement in American Society*. New York: Academic Press.
- Featherman, D. L. and R. M. Hauser
 1976 "Sexual inequalities and socioeconomic achievement in the U.S., 1962-1973." *American Sociological Review* 41:462-83.
- Goldberger, A. S.
 1964 *Econometric Theory*. New York: Wiley.
- Hauser, R. M.
 1971 *Socioeconomic Background and Educational Performance*. Washington, D.C.: American Sociological Association Monograph.
- Hauser, R. M., W. H. Sewell and D. F. Alwin
 1976 "High school effects on achievement." Pp. 309-41 in W. H. Sewell, R. M. Hauser and D. L. Featherman (eds.), *Schooling and Achievement in American Society*. New York: Academic Press.
- Heise, D. R.
 1972 "Employing nominal variables, induced variables, and block variables in path analyses." *Sociological Methods and Research* 1:147-73.
- Hout, M.
 1977 "A cautionary note on the use of two-stage least squares." *Sociological Methods and Research* 5:335-46.
- Hout, M. and W. R. Morgan
 1975 "Race and sex variations in the causes of the expected attainments of high school seniors." *American Journal of Sociology* 81:364-94.

- Jencks, C. S. and M. D. Brown
1975 "Effects of high schools on their students." *Harvard Educational Review* 45:273-324.
- Johnston, J.
1972 *Econometric Methods*. 2nd ed. New York: McGraw-Hill.
- Kerckhoff, A. C.
1974 *Ambition and Attainment*. Washington, D.C.: American Sociological Association Monograph.
- Kerckhoff, A. C. and R. T. Campbell
1977 "Black-white differences in the educational attainment process." *Sociology of Education* 50:15-27.
- McClendon, M. J.
1976 "The occupational status attainment processes of males and females." *American Sociological Review* 41:52-64.
- Marini, M. M.
1974 "Individual and contextual influences on the future orientations of adolescents." Ph.D. dissertation, Department of Social Relations, Johns Hopkins University.
1977a "Explaining the sex difference in educational attainment." Working paper, Population Study Center, Battelle Human Affairs Research Centers.
1977b "Sex differences in the transition to adulthood." Working paper, Population Study Center, Battelle Human Affairs Research Centers.
- Marini, M. M. and E. Greenberger
1978 "Sex differences in educational aspirations and expectations." *American Educational Research Journal* 15:67-79.
- Marini, M. M., A. R. Olsen and D. B. Rubin
1977 "Maximum likelihood estimation in panel studies with missing data." Working paper, Population Study Center, Battelle Human Affairs Research Centers.
- Ornstein, M. D.
1976 *Entry into the American Labor Force*. New York: Academic Press.
- Rosen, B. C. and C. S. Aneshensel
1978 "Sex differences in the educational-occupational expectation process." *Social Forces*. In press.
- Rossi, P. H., W. A. Sampson, C. E. Bose, G. Jasso and J. Passel
1974 "Measuring household social standing." *Social Science Research* 3:169-90.
- Ryder, N. B. and C. F. Westoff
1971 *Reproduction in the United States 1965*. Princeton: Princeton University Press.
- Scanlon, J. H.
1975 *Sex Roles, Life Styles, and Childbearing: Changing Patterns in Marriage and the Family*. New York: Free Press.
- Sewell, W. H. and R. M. Hauser
1977 *Education, Occupation and Earnings*. New York: Academic Press.
- Sewell, W. H., R. M. Hauser and W. C. Wolf
1977 "Sex, schooling, and occupational careers." Working paper 77-31, Center for Demography and Ecology, University of Wisconsin.
- Sewell, W. H. and V. P. Shah
1967 "Socioeconomic status, intelligence, and the attainment of higher education." *Sociology of Education* 40:1-23.
1968a "Parents' education and children's educational aspirations and achievements." *American Sociological Review* 33:191-209.
1968b "Social class, parental encouragement, and educational aspirations." *American Journal of Sociology* 73:559-72.
- Sewell, W. H., A. O. Haller and G. W. Ohlendorf
1970 "The educational and early occupational attainment process: replication and revision." *American Sociological Review* 35:1014-27.
- Spaeth, J. L.
1977 "Differences in the occupational achievement process between male and female college graduates." *Sociology of Education* 50:206-17.
- Taubman, P. and T. Wales
1974 *Higher Education and Earnings*. New York: McGraw-Hill.
- Temme, L. V.
1973 "Prestige scores for 1970 census occupational categories." Unpublished.
- Treiman, D. J. and K. Terrell
1975 "Sex and the process of status attainment: a comparison of working women and men." *American Sociological Review* 40:174-200.
- Turner, R. H.
1966 "On neighborhood context and college plans (I)." *American Sociological Review* 31:698-702.
- U.S. Bureau of the Census
1969 "Marriage, fertility, and childspacing: June 1965." *Current Population Reports, Series P-20, No. 186*. Washington, D.C.: U.S. Government Printing Office.
1973 *United States Census of Population: 1970. Age at First Marriage. Report PC(2)-4D*. Washington, D.C.: U.S. Government Printing Office.
1976a "Educational attainment in the United States: March 1975." *Current Population Reports, Series P-20, No. 295*. Washington, D.C.: U.S. Government Printing Office.
1976b "Number, timing, and duration of marriages and divorces in the United States: June 1975." *Current Population Reports, Series P-20, No. 297*. Washington, D.C.: U.S. Government Printing Office.
- Voss, P. R.
1977 "Social determinants of age at first marriage in the United States." Paper presented at the annual meeting of the Population Association of America, St. Louis.
- Waite, L. J. and K. A. Moore
1978 "The impact of an early first birth on young women's educational attainment." *Social Forces* 57:845-65.
- Wilson, K. L. and A. Portes
1975 "The educational attainment process: results from a national sample." *American Journal of Sociology* 81:343-63.

INSTITUTIONALIZATION AND SICK ROLE IDENTIFICATION AMONG THE ELDERLY*

JOHN F. MYLES

Carleton University

American Sociological Review 1978, Vol. 43 (August):508-521

The purpose of this paper is to identify the effects of institutionalization on patterns of sick role identification among the aged. The paper is an effort to test empirically the implications of Goffman's (1961) general analysis of total institutions and Freidson's (1970) analysis of the social organization of illness within such contexts. These analyses argue that the institutional organization of being ill tends to reinforce patterns of illness behavior among clients by constraining the individual increasingly to adopt attitudes and behavior consistent with the assumption of the sick role. With data drawn from the 1971 Aging in Manitoba study (n = 4,805), an empirical test of this theory is presented by contrasting patterns of illness responses among the institutionalized and noninstitutionalized elderly. Specifically, the extent to which institutionalization alters the association between respondents' objective health status and respondents' self-definition of health status is analyzed. Contrary to the expectations of the theory, the results indicate that, given comparable levels of illness and disability, the institutionalized elderly are less likely to incorporate the illness label into their definition of self. In a subsequent section, the implications of these findings for the individual's overall level of life satisfaction also are identified and discussed.

Since the appearance of Goffman's *Asylums* in 1961, the concept of the "total institution" has occupied a prominent place in the sociological lexicon. The term is a generic one intended to include such diverse organizations as prisons and concentration camps as well as mental hospitals and homes for the aged and the deaf. It is defined as:

... a place of residence and work where a large number of like-situated individuals, cut off from the wider society for an appreciable period of time, together lead an enclosed, formally administered round of life. (Goffman, 1961:xiii)

While Goffman's (1961:xiii) own research was conducted primarily in mental hospitals, his object is to understand "total institutions in general." The chief concern is to identify the manner in which total institutions affect "the structure of the self" (Goffman, 1961:xiii).

In Goffman's view the consequences of institutionalization are devastating. The inmate career is characterized by an inevitable process of "abasements, degradations, humiliations and profanations of self" which are paralleled in the inmate's self-concept by a sense of "personal inefficacy," "failure," "self-pity" and a tendency towards "situational withdrawal" (Goffman, 1961:14, 41, 61, 67). The major mechanism by which this is accomplished is through the imposition of the "inmate role." In Goffman's language, this involves a process of "stripping" in which all other statuses and roles are subordinated to the distinguishing feature of the client which brought him to the institution in the first place. The individual passes through a series of degradation ceremonies, the purpose of which is to strip him of his former identity and provide him with a new identity appropriate to his institutional status. If the organization is to function efficiently, it is necessary that the client recognize and identify with the specific deviant condition being treated by the institution quite independently of the client's actual condition. Such identification will ensure the client's active cooperation in the treatment process.

* This investigation was supported in part by the Division of Research, Planning and Program Development, Manitoba Department of Health and Social Development, Winnipeg, Manitoba, and the Canada Council. I am particularly indebted to Aage B. Sørensen, Michael Aiken, James Greenley and Vivian Wood for their assistance and comments on the research of which this is a part.

At the heart of the general model of total institutions, then, is the systematic manipulation of the client's self-identity, the social reconstruction of the self which is inherent in the process of becoming an inmate. In the context of medical institutions, this implies adoption of a definition of self which is conditioned by the patient or sick role. Inmates may attempt to reject such a label but it is in the interests of the staff to discredit such denials for, if they are to obtain his cooperation, the patient "... must 'insightfully' come to take, or affect to take, the hospital's view of himself" (Goffman, 1961:155).

The rationale for such a process is one that is common to all organizations. As Perrow (1967:197) points out:

Organizations uniformly seek to standardize their raw material in order to minimize exceptional situations. This is the point of de-individualizing processes found in military academies, monasteries or prisons or the superiority of the synthetic shoe material Corfam over leather.

Indeed, in Freidson's (1970) view it is precisely for this purpose that institutions for the ill and disabled exist. He writes:

... when the sick person is institutionalized, the experience of being ill becomes far more amenable to organization by staff demands, for the person tends to lose his social and physical mobility, to be isolated from his lay associates, to be cut off from the information he would need in order to assume an active role in the management of his illness and to be fitted into administrative routines organized to permit the staff to work in ways they consider effective and convenient. In fact only when he is institutionalized can the sick person be restricted to performing only one role—that of the patient. (1970:326)

It is important to recognize that such an analysis is essentially an extension of the labelling or societal-reaction approach to deviance to the study of illness and illness behavior. Within this perspective what is important

... is the *imputation* of deviance to one individual by significant others (including himself). ... Whether or not he is "really" deviant [e.g., sick] or has personal qualities motivating him to be deviant is quite beside the point. What is to the point is the character of the system of social control ...

[which] motivate[s] him to regard himself as deviant. (Freidson, 1970:219)

The total institution represents one of the more extreme forms of the societal reaction to illness. Once the individual is institutionalized "... he can do no more than play his deviant role, choosing only to play it badly or well" (Freidson, 1970:325). Whether or not the individual is really sick becomes irrelevant; what is significant is that inmates inevitably come to believe that they are sick and to behave accordingly. At a social-psychological level, therefore, the institutional organization of being ill is inherently iatrogenic; that is, the treatment becomes the cause of the "dis-ease."

If such a view is correct, then the institutional organization of care for the elderly must be seen as particularly problematic. As numerous studies have shown (Palmore and Luikart, 1972; Spreitzer and Snyder, 1974), the extent to which individuals believe and define themselves to be ill has been found repeatedly to be the single best predictor of life satisfaction and morale among the elderly. Moreover, it is clear that institutions for the aged are organized according to the medical model described by Goffman and Freidson. For example, in the 92 institutions for the aged included in this study, approximately 50% of all staff employed in 1971 were classified as being in medical occupations. Even the provision of personal care services which make up the bulk of the treatment provided by institutions for the aged are cloaked in the medical symbol of the practical nurse. Lines of authority among floor staff are identical to those of the hierarchy of medical training. Although a nonmedical administrator may be the "court of last appeal" for clients, the daily routine of inmates is usually under the direct jurisdiction of the nursing staff. Inevitably, clients come to be identified as patients. They may vary in the degree of disability or illness from which they suffer, but they are patients nonetheless.¹

¹ The singular importance of the medical model in the organization of care for the elderly inevitably was impressed upon me during visits to institutions for the aged in Manitoba during the summer of 1974.

At issue, then, is whether the institution exacerbates or mitigates the social-psychological consequences of illness and disability. As Freidson (1970:206) suggests, "... medicine creates the social possibilities of acting sick . . ." and the institutional organization of being ill ensures that the patient will do so, independently of the actual state of his health. Our task is to determine whether such is indeed the case.

In order to address this question empirically, it is useful to represent the argument mathematically. Following Hernes (1976) terminology, at issue are the parameter values of the process structure which link a set of inputs (the objective health status of an individual) to a set of outputs (the individual's subjective experience of illness). In terms of the usual regression equations we can represent this effect by writing two equations, one for the institutionalized and one for the non-institutionalized, as in:

$$Y_I = a_I + b_1X + e; \quad (1)$$

$$Y_{NI} = a_{NI} + b_{NI}X + e; \quad (2)$$

where Y represents the individual's level of identification with the sick role, X represents a measure of his objective health status and I and NI stand for the institutionalized and noninstitutionalized, respectively. On the basis of the arguments presented by Goffman and Freidson, we would expect that, given comparable levels of illness and disability, the equation for the institutionalized will, in general, indicate higher expected values on Y than the equation for the noninstitutionalized.

For purposes of clarification it is instructive to rewrite equations (1) and (2) as a single equation, as in:

Members of the social work staff of these institutions never failed to note the necessity of undermining the definition of clients as "patients" if they were to gain influence in the running of the organization. Thus far, they had met with little success. Even administrators who wished to demedicalize the institution—e.g., by having nurses dress in street clothes—had been frustrated by the resistance of medical staff whose influence was in no small measure a reflection of the short supply of trained medical personnel who are required by law in all institutions and on whom the daily functioning of the institution is dependent.

$$Y = a + b_1I + b_2X + b_3I \cdot X + e, \quad (3)$$

where I represents a dummy variable coded one if the respondent is institutionalized and zero otherwise. In equation (3), b_1 represents the difference in intercepts between the two populations and b_3 is an interaction term representing the difference in slopes. Given the verbal formulation presented by Goffman and Freidson, sufficient evidence to support the theory would be provided by a significant positive coefficient for b_1 if $b_3 = 0$. Such a finding would indicate that, given comparable levels of illness/disability, the institutionalized elderly are more likely to incorporate the illness label into their definition of self than the noninstitutionalized. As usually presented, the argument makes no assumptions about the value of b_3 ; i.e., there is no explicit statement that the effect of institutionalization on the relationship between X and Y will vary with the individual's level of disability. However, in order to assess the generalizability of the theory such an eventuality must clearly be taken into account.

DATA AND METHODS

The analysis is based on data gathered by the Department of Health and Social Development of the Province of Manitoba in 1971 on a stratified random sample of the population aged 65 and over of the Province of Manitoba. The sample ($n = 4,805$), which was stratified by region (Metro-Winnipeg/Non-Metro) and residential status (institutionalized/noninstitutionalized), was selected from the registry of the Manitoba Health Services Commission. A detailed discussion of the sampling design and sample characteristics can be found in *Aging in Manitoba* (Province of Manitoba, 1973).

In instances where the respondent was too ill to answer the questionnaire a proxy was employed. Four percent of the interviews in the noninstitutionalized population and 24.9% in the institutionalized population were conducted in this manner. Since numerous self-report items and all attitudinal items were excluded from the proxy interviews, the present analysis is restricted to that portion of the popula-

tion functionally capable of responding to the interview. In addition, all respondents who were not living in the general community or in one of the 92 nursing homes or hostels for the aged in the province were excluded from the analysis (e.g., residents of mental institutions, housing units for the aged). As a result, the analysis presented below is based on a sample of 473 institutionalized and 3,378 noninstitutionalized respondents.²

The data analysis relies principally on the use of partial regression coefficients. However, because disproportionate stratified sampling was employed, the use of significance tests based on assumptions of simple random sampling becomes problematic. On the one hand, the use of stratification increases the efficiency of the design and hence one is in danger of making a Type II error when using tests which assume simple random sampling. In contrast, in order to obtain population estimates it is necessary to weight the data, thereby artificially increasing the sample size and hence introducing the danger of making a Type I error. In practice it was found that the effect of weighting was small. Replication of the analysis indicated that when the metric coefficients produced by the weighted analysis were

more than twice their standard error, then the coefficients of the unweighted analysis were significant at the .05 level. Moreover, given the sample size, most coefficients large enough to be of any substantive interest were much larger than twice their standard error.

In the presentation of results all coefficients which were less than twice their standard error were set equal to zero and the equations reestimated. Hence all non-zero coefficients presented in the following tables can be assumed to be greater than twice their standard error. For clarity of presentation the results are presented in separate equations for the two comparison groups. However, the metric coefficients are derived from a single equation of the form presented in equation (3) above. While the results are identical in either method, the use of a single equation has the advantage that it enables one to test for significant differences in parameter values.

In order to test the argument empirically we require a measure of the extent to which different individuals perceive and evaluate the state of their health. In the Manitoba study, each respondent was asked the following question:

For your age, would you say, in general, your health is:

1. Excellent—Never prevents activities;
2. Good for your age—Rarely prevents activities;
3. Fair for your age—Occasionally prevents some activities;
4. Poor for your age—Very often prevents many activities;
5. Bad for your age—Health troubles or infirmity all the time—Prevents most activities.

Such a measure is appropriate to our task for a number of reasons. First of all, it identifies the tendency of respondents to adopt or reject the illness label not in the abstract, but rather against a norm for their own age cohort and thereby indicates the degree to which an individual views himself as deviant. Further, the question is specified by the extent to which respondents view the state of their health as problematic or disruptive with respect to continued social functioning. In effect, the

² Unpublished data from a study conducted in 1973 by Statistics Canada indicate that institutional care in Manitoba is quite similar to that found elsewhere in Canada with respect to such factors as percentage of the elderly who are institutionalized, percentage of proprietary (profit-making) institutions and level of care provided. Canadian institutions in general and those in Manitoba in particular tend to be subject to much tighter regulation by the state (provincial governments) than is true in the U.S., where both the level and enforcement of legislation have been defined as major problems (Mendelson, 1974). There are also proportionately fewer proprietary institutions in Manitoba and Canada generally than is the case in the U.S. In the absence of comparative data on the "aging industry" in the two countries, making comparative judgements is a hazardous enterprise. However, the opinion of informed observers is that many of the abuses which characterize the American nursing home industry (U.S. Senate Committee on Aging, 1974) are less prevalent in Canada. Such differences, however, are irrelevant for a test of the theory since it is not based on assumptions about quality of care but rather on the existence of bureaucratically organized hierarchical relations between staff and clients (Goffman, 1961:6) which are characteristic of both systems.

measure identifies the degree to which the respondent perceives the state of his health to be congruent or discrepant with what he might normally expect for his age as measured by his level of social functioning. As such, it constitutes an indicator of the individual's incorporation of the illness label into his definition of self.³

Table 1 presents the distribution of self-assessed health status within the two comparison groups. The differences which appear between the two groups are extremely small but might be construed initially as giving some support to the institutionalization thesis. The institutionalized population has a slightly larger percentage reporting that their health is either poor or bad for their age. On reflection, however, the small differences which appear in Table 1 are somewhat surprising in view of the much different levels of illness and disability which characterize the institutionalized and noninstitutionalized populations (see Table 3). It is probable that the slightly greater tendency of the institutionalized to define themselves as ill is in fact a reflection of the true situation and that, once the respondent's actual health status is controlled, a quite different picture will emerge. In effect, a multivariate strategy as outlined above is required to test the theory. Of particular importance in such an analysis is the identification and measurement of the respondent's actual (objective) health status.

The concepts of objective health and disease are notoriously difficult ones to define and operationalize (see Haberman, 1969; Hennes, 1972). For purposes of the present analysis, objective health was defined operationally as the absence of physical impairment and disease and included the following indicators:

(i) Functional Incapacity. The first of these measures is a modified version of Shanas et al.'s (1968) Index of Incapacity, which measures the individual's ability to move about freely, maintain personal hygiene, etc. The scale is made up of thirteen activities of daily living on which in-

Table 1. Self-Assessed Health Status by Residential Status, Population 65+, Manitoba, 1971¹

	Noninstitutionalized	Institutionalized
1. Excellent	14.9%	11.6%
2. Good for your age	48.5	46.3
3. Fair for your age	27.6	29.8
4. Poor for your age	8.2	10.5
5. Bad for your age	0.9	1.8

¹ Noninstitutionalized (n=3,370), n*=4,651; Institutionalized (n=471), (n*=707), where n=unweighted sample and n*=weighted sample.

dividuals were asked to indicate whether they either received or required assistance.

(ii) Number of Chronic Illnesses. From a list of fifteen chronic illnesses characteristic of the elderly, respondents were asked to indicate those from which they suffered (e.g., arthritis, rheumatism, diabetes, heart and circulatory problems, etc.). Each respondent was then assigned a total score indicating the number of chronic conditions from which he or she suffered.

(iii) Hospitalization and Confinement to Bed. Since the preceding measure does not identify the seriousness of a given condition, respondents were also asked to indicate (a) the length of time they had spent in hospital during the past year, and (b) the amount of time they had been confined to bed due to illness.⁴

These three categories of measures correspond exactly to the major conceptual categories which Hennes and Wharton (1970) found were used by respondents as a basis for judging their health—namely, activities, clinical conditions, and medical care system criteria.

The use of multiple measures raises the issue of whether each is simply an indicator of a single underlying construct—illness—or whether in fact we are measur-

³ For a concise review of self-assessed health status among the elderly see Riley and Foner (1968:292-4).

⁴ Both of these measures were scaled ordinally as follows: 0 = none; 1 = less than one week; 2 = one week to one month; 3 = more than one month.

ing several distinct phenomena. For policy purposes, it is frequently desirable to keep such phenomena as chronic illnesses and functional incapacity distinct. Institutions for the aged can in general do little to cure chronic illnesses, but potentially can do much to alleviate the consequences of functional disability. Moreover, not all illnesses lead to disability as indicated by the relatively low zero-order correlation (.20) between number of chronic illnesses and functional disability for the population as a whole. On the other hand, our introduction of "time spent in bed" and "time spent in hospital" during the past year is based primarily on a recognition that our measure of chronic illness does not take into account the seriousness of the disease conditions.

In addition to the conceptual problem, there is an empirical problem which arises from using multiple measures of health. In testing for interaction effects between institutionalization and objective health status, we are spreading the effects of institutionalization across four different independent variables and thereby arbitrarily reducing the possibility of finding significant interactions.

In order to determine whether the four health indicators were measuring one or several dimensions of physical health, they were analyzed by means of a classical (R-type) factor analysis. The results of this analysis produced one underlying factor with eigenvalue exceeding unity which accounted for approximately 40% of the total covariance among the four variables. The factor weights derived from the least-squares estimate for the factor (Smith, 1974) were then used to construct a composite index of physical disability (PDI). This weights the standardized values of the variables by the partial regression coefficients obtained by regressing the observed variables on the inferred underlying factor before summing them to form the index. A high score of this index is indicative of high disability and a low score indicative of low disability. The factor loadings and factor weights are presented in Table 2.

The use of a single measure of health status has the advantage that it increases the probability that we shall be able to

Table 2. Factor Loadings and Factor Weights for Four Measures of Objective Health Status

	Factor Loadings	Factor Weights
Functional Incapacity	.39	.23
Number of Chronic Illnesses	.45	.28
Time Spent in Hospital	.50	.33
Time in Bed Due to Illness	.44	.27

detect interaction effects (slope differences) if such effects are present. However, since the technology available to institutions for the aged is oriented primarily towards minimizing the consequences of functional incapacity and can do little with respect to relieving the effects of chronic illness, we have a substantive interest in examining each of the four components separately. Accordingly, in identifying the manner in which institutionalization alters the relationship between perceived health status and objective health status, we shall present results for both the composite index and its separate components.

Other background characteristics which were included in earlier stages of the analysis included age, sex, marital status, income, education, etc. Of these, only age proved to have a significant effect on self-assessed health status and hence the remaining control variables are excluded from the present analysis.

FINDINGS

Table 3 presents the mean scores for each of the variables included in the analysis for the two groups as well as the zero-order correlations of the independent variables with self-assessed health status. As indicated by the positive correlation coefficients, all of the measures of illness and disability are associated with a tendency to define oneself as ill. As might be expected, the institutionalized score higher on all measures of disability and illness, with the greatest difference being found with respect to functional incapacity.

Table 4 presents the results of the multivariate analysis for the two populations. The standardized coefficients presented in

Table 3. Mean Values and Zero-Order Correlations of Health Status Measures and Age with Self-Assessed Health Status, Population 65+, Manitoba, 1971

	Means		Zero-Order Correlations with Self-Assessed Health	
	Non- institutionalized	Institu- tionalized	Non- institutionalized	Institu- tionalized
Self-Assessed Health	2.32	2.45		
Physical Disability Index	-.029	.664	.46	.43
Functional Incapacity	.41	3.81	.27	.24
Number of Chronic Illnesses	3.57	4.05	.48	.35
Time in Hospital	1.43	1.72	.22	.20
Time in Bed Due to Illness	1.29	1.33	.24	.29
Age	73.60	82.91	.04	-.07

panels 1 and 2 indicate a comparable rank ordering of the relative importance of each of the independent variables, with the exception of confinement to bed due to illness which acquires a somewhat greater importance within the institutionalized population.

In order to identify the manner in which institutionalization alters the relationship between objective and perceived health status, it is necessary to examine the metric coefficients and intercept terms for the two populations presented in panels 3 and 4 of Table 4. Briefly summarized, both

sets of equations indicate that given comparable levels of illness and disability, the institutionalized are less likely than noninstitutionalized respondents to view themselves as ill. Such findings are precisely the opposite of those expected on the basis of the analyses of Goffman and Freidson.

When the Physical Disability Index is used as the measure of health status (panel 4), there is both a significant main effect of institutionalization (as measured by the difference in intercepts, .11) and a significant interaction (as measured by the dif-

Table 4. Regression of Self-Assessed Health Status on Selected Variables by Residential Status, Population 65+, Manitoba, 1971¹

Independent Variables								
	PDI ^a	FUNCT ^a	SICK ^a	HOSP ^a	BED ^a	AGE	a	R ^a
Standardized Coefficients								
1(a) Noninstitutionalized		.16	.41	.08	.12	-.06		.28
1(b) Institutionalized		.17	.28	.10	.19	-.07		.22
2(a) Noninstitutionalized	.47					-.04		.22
2(b) Institutionalized	.43					-.07		.19
Metric Coefficients								
3(a) Noninstitutionalized		.10	.25	.07	.15	-.008	1.69	
3(b) Institutionalized		.04	.25	.07	.15	-.008	1.69	
4(a) Noninstitutionalized	.59					-.005	2.68	
4(b) Institutionalized	.42					-.005	2.57	

¹ Noninstitutionalized (n=3,347, n*=4,619); Institutionalized (n=462, n*=693), where n=unweighted sample, and n*=weighted sample.

^a Physical incapacity index.

^a Functional incapacity.

^a Number of chronic illness.

^a Time spent in hospital during past year.

^a Time spent in bed during past year.

ference in slopes, .17).⁵ Though small, the direction of the differences is the opposite of that anticipated by the theory. Given comparable levels of disability, the institutionalized elderly are less likely to view themselves as disabled. For those who have little disability there is a moderate reduction in sick role identification, but this reduction increases as disability increases.

Panel 3 of Table 4 provides the equations estimated when all four measures of objective health status were entered as independent variables. Using such a procedure reduces the probability of finding significant effects of institutionalization since these effects are now being distributed across five different parameters. The only significant effect of institutionalization found in equations (3:a) and (3:b) is associated with the coefficients for functional incapacity. The effect of functional incapacity on self-assessed health status among the institutionalized is less than half the size of the effect found among the noninstitutionalized (.04 vs. .10). The fact that the effect of institutionalization is mediated primarily through functional incapacity is indicative of the kind of technology available to institutions for the aged, a technology which is primarily oriented towards minimizing the consequences of disability and, in general, can do little with respect to the typical chronic diseases from which the elderly suffer.

In brief, then, the findings of the present analysis not only do not support the theory, but also provide evidence which suggests the contrary view. Institutions for the elderly are designed to provide a protective, prosthetic environment for the elderly individual with declining physical capacity. The findings presented above suggest that such institutions are also suc-

cessful in providing subjective relief from the experience of illness, particularly for the most disabled. The implications of these findings for our understanding of institutionalization will be discussed in more detail below.

Before doing so, however, it will be useful to examine some alternative interpretations of the findings. The nature of the findings would seem to exclude the possibility that they can be accounted for by self-selection. To the extent that self-selection is operative, we would expect it to produce results opposite to those found in the analysis; that is, given comparable levels of illness and disability, we would expect that those individuals entering an institution would also be more likely to define themselves as being in poor health.

A more plausible competing interpretation, however, is that our results are simply a reflection of a reference group effect. By virtue of becoming institutionalized, individuals are separated from the larger community and brought into contact with individuals who suffer from comparable levels of illness and disability. Accordingly, it is plausible to assume that institutionalization is altering the standards or criteria against which the individual is assessing his health status. Rather than minimizing the disruptive consequences of illness and disability, it may be that, within the institution, these consequences are simply more likely to be accepted as being normal. In other words, our findings do not necessarily indicate that the institutionalized elderly are responding to the supportive services and physical environment of the institution since it is plausible to assume that a similar result could be produced simply by virtue of becoming a member of a community of the sick.

In order to test for this possibility, the relationship between objective health status and self-assessed health status was estimated, controlling for the level of illness and disability within each institution.⁶ If the reference group effect were

⁵ Since the Physical Disability Index does not have an absolute zero point, the difference in intercepts is not particularly meaningful. What is crucial is the value of the index at which the regression lines for the two populations intersect. Since, in fact, the intersection falls outside the range of the observed values of the index for the institutionalized population, the interpretation of the parameters provided in the text is unaffected by this property of the scale.

⁶ The contextual measure was the percentage of clients receiving medium and heavy nursing care within each institution using provincial

operative we would expect to find that as the average level of illness and disability across institutions increases, the tendency of inmates to view themselves as ill will decline. Inmates in institutions where the typical level of disability is high should be more likely to perceive their own level of disability as being normal and hence less likely to define themselves as ill.

When the contextual measure representing the level of disability within each institution was introduced to the equation including all four health measures, it did indeed have a significant effect on self-assessed health status ($Beta = .09$). However, the sign of the coefficient was opposite to that expected on the basis of the preceding discussion. Being in the company of others who were sick/disabled appeared to reinforce the tendency of inmates to define themselves as sick. However, when the contextual measure was introduced to the equation in which health was measured by the Physical Disability Index, this effect disappeared. This latter finding suggests that the effect found previously was simply reflecting measurement error eliminated by the single index of disability. Given that clients are recruited into different institutions on the basis of their level of disability, the variation captured by the contextual measure is likely to reflect individual differences rather than any contextual effect (Hauser, 1970).

In sum, our analysis provides no support for the hypothesis that the effects of institutionalization found earlier can be explained in terms of a reference group effect. Given comparable levels of illness and disability, the institutionalized are less likely to define themselves as ill and there is no evidence that they are simply responding to a social milieu in which a high level of illness and disability is considered normal. We are left with the conclusion that institutions for the aged do provide their elderly clients with a prosthetic environment which is reflected in

the inmate's subjective experience of institutional life. Relief from illness, then, constitutes one dimension of the manner in which institutionalization affects the structure of the self among the aged.

INSTITUTIONALIZATION, HEALTH, AND LIFE SATISFACTION

A considerable body of literature on the elderly points to the singular importance of health and illness in determining life satisfaction and morale among the elderly (Maddox and Eisdorfer, 1962; Streib, 1956; Kutner, 1956). As already noted, multivariate analyses (Palmore and Luikart, 1972; Spreitzer and Snyder, 1974) have found self-assessed health status to be the single best predictor of life satisfaction among the elderly. Spreitzer and Snyder's (1974) analysis also indicates that this association increases with age (i.e., when comparisons between those over and under 65 are made). Hence, the analysis of the effects of institutionalization on self-assessed health status clearly warrants our attention. By increasing the inmate's evaluation of his health status, the institution is by implication also increasing the respondent's overall level of satisfaction. Since the client is less likely to find his illness and disabilities disruptive, we would assume that his level of life satisfaction will be enhanced. Is this assumption warranted?

To sustain this interpretation of our findings, it is necessary to show that the effects of actual and perceived disability on life satisfaction among the institutionalized are either identical to, or less than, the effects found among the noninstitutionalized. If the effect is larger, then we would have to reassess our earlier interpretation. If life satisfaction among the institutionalized declines more rapidly as disability increases, then the two effects may cancel each other out and we would have to conclude that given comparable levels of disability, life satisfaction will be the same for the two groups.

The measure of life satisfaction employed is Neugarten et al.'s (1961) Life Satisfaction Index (A), one of the more commonly used measures of satisfaction in studies of the aged. The means and

classifications. Medium nursing care indicates a patient requires assistance with 50% of all activities of daily living (feeding, bathing, dressing, etc.) and heavy nursing care indicates a patient requires assistance with 75% of such activities and/or is bedridden.

Table 5. Mean and Standard Deviation of Life Satisfaction by Residential Status, Population 65+, Manitoba, 1971¹

Population	Mean	Standard Deviation
Noninstitutionalized	13.64	3.75
Institutionalized	12.65	4.05

¹ Noninstitutionalized (n=3,299; n*=4,553); Institutionalized (n=433, n*=649), where n=unweighted sample and n*=weighted sample.

standard deviations of life satisfaction in the two populations are presented in Table 5. The mean level of satisfaction among the institutionalized is approximately one point lower than that found among the noninstitutionalized.

Table 6 presents the zero-order correlations and the standardized and metric coefficients from the regression of life satisfaction on the measures of subjective and objective health status. Once functional disability and number of chronic illnesses are controlled, there is no

additional effect of time spent in hospital or time spent in bed and hence these two measures are excluded from the analysis. As in the previous section, however, we shall present separate results for the combined index of physical disability and for its separate components. The coefficients are derived from equations which included all other factors found to affect life satisfaction, namely, measures of the respondent's economic status, marital status, and level of social participation.⁷

As indicated by the zero-order correlations (panel 1) there are significant associations between life satisfaction and both the subjective and objective health measures in both populations. The strongest association is with perceived health status

⁷ The results for the full equations can be found in Myles (1977). The general findings were highly consistent with previous research on life satisfaction and aging (Adams, 1971), with respect to both the variables included in the model and the relative importance of each in accounting for explained variation in life satisfaction.

Table 6. Regression of Life Satisfaction on Subjective and Objective Health Status, by Residential Status, Population 65+, Manitoba, 1971¹

Population	Independent Variables			
	SAH ^a	PDI ^b	FUNCT ^c	SICK ^d
Zero-Order Correlations with Life Satisfaction				
1(a) Noninstitutionalized	-.36	-.27	-.19	-.27
1(b) Institutionalized	-.34	-.22	-.19	-.22
Standardized Coefficients				
2(a) Noninstitutionalized	-.23		-.08	-.10
2(b) Institutionalized	-.21		-.09	-.11
3(a) Noninstitutionalized	-.27	-.12		
3(b) Institutionalized	-.23	-.10		
Metric Coefficients				
4(a) Noninstitutionalized	-1.01		-.27	-.25
4(b) Institutionalized	-1.01		-.10	-.25
5(a) Noninstitutionalized	-1.18	-.68		
5(b) Institutionalized	-1.18	-.23		

¹ Coefficients are estimated for equations, which control for economic status, social participation, and marital status. R² for the institutionalized population=.22; R² for noninstitutionalized=.24. Intercept terms for the two populations were not significantly different.

^a Noninstitutionalized (n=3,199, n*=4,415); Institutionalized (n=433, n*=649), where n=unweighted sample, n*=weighted sample.

^b SAH=Self-Assessed Health Status.

^c PDI=Physical Disability Index.

^d FUNCT=Functional Incapacity.

^e SICK=Number of Chronic Illnesses.

and, as indicated by the standardized coefficients in panels 2 and 3, the effects of objective health on life satisfaction, for the most part, are mediated through this variable. Consistent with the literature cited above, self-assessed health status is the single most important determinant of life satisfaction for both the institutionalized and noninstitutionalized populations as measured by the standardized regression coefficients for all variables which predict life satisfaction. In order to identify whether and to what extent institutionalization alters the relationship between health status and life satisfaction, we must turn to the metric coefficients presented in panels 4 and 5.

The first finding worthy of note is our failure to find any significant difference in the effects of self-assessed health status on life satisfaction across the two populations. The effects, as measured by the metric coefficients, are uniformly large and identical. A one point change in self-assessed health status leads to slightly more than a one point change in life satisfaction in both groups. This finding, of course, is important with respect to the interpretation of the results presented in the previous section. Respondents in both populations are affected equally by their subjective perceptions of health status, but institutionalization reduces the probability that individuals will evaluate their health as poor. Hence, institutionalization indirectly increases the level of life satisfaction among residents of institutions for the aged by lowering their level of perceived disability.

Among the institutionalized and noninstitutionalized populations, there are also significant residual effects of objective health status on life satisfaction which are not accounted for by the respondent's own perceptions of his or her health status. This is simply indicative of the fact that illness and disability may generate problems for individuals without them being directly aware of the source of their difficulties. As indicated by the coefficients for the Physical Disability Index (panels [5:a] and [5:b]), the effect of disability on life satisfaction among the institutionalized is approximately one-third the size of the effect found among the

noninstitutionalized population. Hence, the additional effect of physical disability on life satisfaction is much less strong among the institutionalized. As indicated in panel 4, where the separate measures of objective health status are used as independent variables, this is primarily due to the manner in which institutionalization alters the relationship between functional incapacity and life satisfaction, a finding which again is consistent with the characterization of the institution as a prosthetic environment.

DISCUSSION

Since our findings generally contradict what was expected on the basis of the analyses of Goffman and Freidson, what does this suggest for our understanding of total institutions?

Preeminent among the factors which lead individuals to define a set of symptoms as deviant is the extent to which such symptoms are disruptive of what they consider to be their normal level of functioning (Mechanic, 1968:145-6). DiCicco and Apple (1960), for example, surveyed a sample of the elderly and concluded that health was defined in terms of activity—to be active was to be healthy. Poor health represented a state of health which interfered with daily activities. Parsons (1958:176) has formalized this view by defining health as the "state of optimum capacity of an individual for the effective performance of the roles and tasks for which he has been socialized."

At the subjective level, then, becoming ill implies recognition that one's physical or mental condition is abnormal or deviant in terms of one's capacity to pursue one's normal or preferred set of life activities. However, the potential disruption which arises from any given condition is determined by the level of physical or mental functioning required within a particular social context. A sore back, for example, will have different implications for a laborer than for an office worker. For the former, it may require sick leave, while for the latter, it is only a minor annoyance.

The purpose of institutions for the aged is to minimize the disruptive potential of physical incapacity and illness. If they are

successful in this effort, we should expect this to be reflected in their clients' subjective experiences of illness. In effect, given comparable levels of illness and disability, we should expect the institutionalized elderly to be less likely to view themselves as sick, ill, or disabled, i.e., to incorporate the illness label into their definition of self. Since, in general, institutions for the aged are unable to "cure" their clients, such an outcome must surely constitute one of the major measures of organizational effectiveness for such agencies.

Precisely because illness is a *social* state, defining oneself as ill is also contingent on having this definition validated by significant others. In many instances, it may be necessary that such validation come from an individual who has been officially appointed to fulfill this function, e.g., a physician. Indeed, the fact that significant others define the subject as ill can be a major factor in leading the subject to define himself as ill. Moreover, just as the subject is likely to view himself as ill when a set of symptoms prove to be disruptive of his normal functioning, so too are others likely to perceive the individual as ill when a set of symptoms leads to forms of behavior which are disruptive of the normal social routine. As Mechanic (1968:145) notes:

... aged persons are hospitalized frequently not so much because they have serious symptoms that can be helped through hospitalization but because their presence in the household is disruptive and inconvenient.

Accordingly, Mechanic (1968:147) argues that an important contributing factor to the adoption of the illness label is the tolerance threshold of those who are exposed to and evaluate the deviant signs and symptoms.

In brief, subjective responses to illness tend to be a function of the level of disruption (dis-ease) which results for both the subject and significant others from a set of symptoms. The question, then, is whether the declining functional capacity characteristic of the institutionalized aged is likely to prove more disruptive within the institution or outside of it. Are agency officials likely to be more or less tolerant than those individuals the client is likely to encounter in the larger society? In sum, is

it the institution or the larger society which is likely to be a more difficult environment for the disabled elderly?

Precisely because it is the job of institutional staff to care for the sick and disabled we would expect their tolerance of the disabled elderly to be greater than that which the individual is likely to encounter elsewhere in society. Freidson (1970:238) himself notes that, in contrast to the layman's view of illness, "the ideology . . . of contemporary professionals in the health field asserts that to the professional all is legitimate, that there is no illegitimate illness." The technology available to institutions for the aged is systematically designed to minimize the disruption, both for the individual and others, which is associated with increasing functional incapacity. This is precisely what is implied by our characterization of the institution as a prosthetic environment. Accordingly, assuming comparable levels of disability and disease, we would expect the institutionalized to be much less likely to view their illnesses and disabilities as problematic or as obstacles to their continued well-being and social functioning. As a result, they should also be less likely to incorporate the illness label into their definitions of self—to view themselves as sick, ill, or disabled. In effect, if institutions for the aged are successful in creating a prosthetic environment, then this should be reflected in the inmate's experience by a reduction in the role which illness plays in shaping the structure of the self.

Sociologists, however, have long been enamoured by paradox, internal contradictions, latent functions and unanticipated consequences. Not surprisingly, then, sociological analyses of the "institutional organization of being ill" (Freidson, 1970:322) have postulated that whereas the intended effects of institutionalizing the physically or mentally disabled may be therapeutic or prosthetic, the actual effects at the social-psychological level are iatrogenic. By virtue of organizing the lives of clients around the management of illness, the institution reinforces the tendency of clients to view themselves as sick, ill, or disabled and to incorporate such labels into their definition of self. In effect, the institutional organization of being

ill is seen as exacerbating rather than relieving the subjective consequences of illness and disability. Our findings have indicated that just the opposite is true.

Our findings, however, should be interpreted relative to the standard of comparison which has been employed in the analysis, namely, the noninstitutionalized elderly who suffer from comparable levels of illness and disability but who generally do not have access to the same level of goods and services provided within the institutional context. Accordingly, the findings are as revealing with respect to the conditions of life experienced by the noninstitutionalized elderly as they are with respect to the institutionalized. Our findings then should not lead to the conclusion that the institutional organization of care for the elderly is the most efficient means of providing care for the elderly but, rather, given the general situation of the elderly within the larger society, the institution provides a less problematic living environment for the elderly individual suffering from declining functional capacity.⁸

Nor should the findings be construed as a refutation of the societal reaction approach to illness and illness behavior. What they do suggest is that in assessing this reaction Goffman and his disciples have perhaps identified the wrong enemy. In general, such analyses have tended to treat the institution in isolation from the wider social context. As a result, the institution is identified as the villain and the larger society which creates the demand for such institutions emerges unscathed. Goffman himself notes this potential source of bias in his analysis when he writes (1961: x):

⁸ An anonymous reviewer has pointed out the danger that these findings might be misinterpreted as indicating a policy of institutionalizing more elderly persons. There are other reasons for rejecting such a solution, not the least important of which is the preference of most elderly persons to remain in the larger community. An equally plausible solution is the provision of the noninstitutionalized elderly with a level of goods and services comparable to that currently available to the institutionalized elderly. All too often, however, arguments for deinstitutionalization have been utilized to legitimate a reduction in the total level of goods and services made available to the elderly population as a whole.

... I want to warn that my view is probably too much that of a middle-class male; perhaps, I suffered vicariously about conditions that lower-class patients handled with little pain.

In effect, Goffman's conclusions are based on the adoption of the position of a (healthy) middle-class male as the standard of comparison for evaluating the content of the inmate career. A more valid standard of comparison, and the one adopted here, would seem to be the position of comparable individuals who must cope with the larger society which in general has little use and less tolerance for such negatively privileged status groups as the aged and mentally ill.

REFERENCES

- Adams, David
1971 "Correlates of satisfaction among the elderly." *The Gerontologist* 11:64-8.
- DiCicco, L., and D. Apple
1960 "Health needs and opinions of older people." Pp. 26-39 in D. Apple (ed.), *Sociological Studies of Health and Sickness*. New York: McGraw-Hill.
- Freidson, E.
1970 *Profession of Medicine*. New York: Dodd Mead.
- Goffman, E.
1961 *Asylums*. Garden City: Doubleday.
- Haberman, P. W.
1969 "The reliability and validity of the data." Pp. 343-83 in John Kosa, Aaron Antonovsky and Irving K. Zola (eds.), *Poverty and Health*. Cambridge, Ma.: Harvard University Press.
- Hauser, R.
1970 "Contexts and context: a cautionary tale." *American Journal of Sociology* 75:645-64.
- Hennes, J. D.
1972 "The measurement of health." *Medical Care Review* 29:1268-88.
- Hennes, J. D. and T. Wharton
1970 *Consumer Criteria for Measures of Health*. Columbia: School of Medicine, University of Missouri.
- Hernes, Gudmund
1976 "Structural change in social processes." *American Journal of Sociology* 82:513-47.
- Kutner, B.
1956 *Five Hundred Over Sixty*. New York: Russell Sage.
- Maddox, G. and C. Eisdorfer
1962 "Some correlates of activity and morale among the elderly." *Social Forces* 40:254-60.
- Mechanic, D.
1968 *Medical Sociology: A Selective View*. New York: Free Press.
- 1972 "Sociological issues in mental health." Pp.

- 203-22 in D. Mechanic, *Public Expectations and Health Care*. New York: Wiley.
- Mendelson, M. A.
1974 *Tender Loving Greed*. New York: Vintage.
- Myles, John F.
1977 *Institutionalizing the Elderly: An Empirical Assessment of the Sociology of Total Institutions*. Ph.D. dissertation, Department of Sociology, University of Wisconsin, Madison.
- Neugarten, B. L., R. J. Havighurst, and S. S. Tobin
1961 "The measurement of life satisfaction." *Journal of Gerontology* 16:134-43.
- Palmore, E. and C. Luikart
1972 "Health and social factors related to life satisfaction." *Journal of Health and Social Behavior* 13:68-80.
- Parsons, Talcott
1958 "Definitions of health and illness in light of American values and social structure." Pp. 165-87 in E. Jaco (ed.), *Patients, Physicians and Illness*. Glencoe: Free Press.
- Perrow, C.
1967 "A framework for the comparative analysis of organizations." *American Sociological Review* 35:503-15.
- Province of Manitoba
1973 *Aging in Manitoba*, Vol. 1. Winnipeg: Manitoba Department of Health and Social Development.
- Riley, M. W. and A. Foner
1968 *Aging and Society*, Vol. 1. New York: Russell Sage.
- Shanas, E., P. Townsend, D. Wedburn, H. Friis, P. Milhoj, and J. Stehouwer
1968 *Old People in Three Industrial Societies*. New York: Atherton.
- Smith, K.
1974 "Forming composite scales and estimating their validity through factor analysis." *Social Forces* 53:168-80.
- Spreitzer, E. and E. E. Snyder
1974 "Correlates of life satisfaction among the aged." *Journal of Gerontology* 29:454-8.
- Streib, G.
1956 "Morale of the retired." *Social Problems* 3:270-6.
- U.S. Senate Committee on Aging
1974 *Nursing Home Care in the United States: Failures in Public Policy*. Washington, D.C.: U.S. Government Printing Office.

STATUS RESOURCES, SOCIETAL REACTIONS, AND TYPE OF MENTAL HOSPITAL ADMISSION*

WILLIAM A. RUSHING

Vanderbilt University

American Sociological Review 1978, Vol. 43 (August):521-533

This paper reports on a test of the status resource hypothesis of type of mental hospitalization (voluntary vs. involuntary). Findings are based on all 21-64 year-old first admissions to all state mental hospitals in Tennessee between 1956 and 1965. Analysis based on the linear probability estimation model (Grizzle et al., 1969) consistently supports the hypothesis. The analysis also supports two interaction hypotheses, namely, that the effect of status resources varies depending on psychiatric status and level of behavioral deviance.

Studies of the relationship between socioeconomic status and type of admission to mental hospitals indicate that patients committed by the courts are more apt to have fewer socioeconomic resources than patients who enter voluntarily (Linsky, 1970; Rushing, 1971; Gove and Fain, 1977; Rushing and Esco, 1977;

Rice, 1977). This is consistent with a status resource hypothesis, which states that individuals with more resources are better able to control their fates and hence to resist legal coercion that would lead to hospitalization (see Rushing, 1971). The results of these studies are also central in debates concerning the relative validity of the labelling perspective and psychiatric perspective on mental illness (Scheff, 1974; Gove, 1975; see also Gove, 1976a; 1976b; Conover, 1976; and Lemert, 1976). Although evidence appears to support the

* I want to thank George Becker, Ivar Berg, Walter Gove, Michael Hughes, and Michael Swafford for their comments on an earlier version and especially Michael Swafford for his consultation concerning the data analysis.

view that social status and not just psychiatric status is an important factor in type of hospitalization, Gove (1975:246) believes that the evidence is as "consistent with the psychiatric perspective as . . . the labelling perspective" because committed patients are apt to be more seriously disturbed (see also Gove and Howell, 1974:88). Indeed, some evidence does show that committed patients may be more disturbed than other patients (Rushing and Esco, 1975; Gove and Fain, 1977). It is not clear, however, that the statistical effect of social status on type of admission is an artifact of individual psychiatric status. The present paper seeks to determine if this is the case by presenting a test of the status resource hypothesis in which the psychiatric status of patients is controlled.

SOCIOLOGICAL RELEVANCE OF THE DISTINCTION BETWEEN VOLUNTARY AND INVOLUNTARY ADMISSIONS

The disadvantages of involuntary admission to mental hospitals are well-known. Committed patients may suffer courtroom trauma and certain indignities not experienced by other patients, and they are deprived of certain civil liberties, such as voting privileges, the right to get a divorce, or to control their economic resources, although the specific deprivations vary from state to state (Ennis and Siegel, 1973). Also, most authorities believe that voluntary hospitalization is preferable on therapeutic grounds (for a review and references, see Gove and Fain, 1977). In addition, studies show that even with level of assessed impairment controlled, the length of stay is longer for those patients who are committed (Rushing and Esco, 1975; Gove and Fain, 1977).

The theoretical significance of the voluntary-involuntary distinction for sociology can be appreciated when it is recognized that types of hospitalization represent different forms of collective reactions to individual behavior which may (or may not) be disruptive to others. Involuntary admissions are formally institutionalized methods communities have for isolating individuals from others. They involve more formalized coercion than do

voluntary admissions. In addition, unless a distinction between voluntary and involuntary admissions is made, certain questions that are crucial to understanding the role of social factors in hospital admissions cannot be answered (indeed, cannot even be raised meaningfully). A review of Gove and Howell's (1974) paper will illustrate this.

Gove and Howell examine the relationship between social resource variables (income and marital status) and severity of illness, and find that resources and severity are related inversely. They conclude that resources facilitate entry into psychiatric treatment and that "when severity of disorder is controlled for, married persons and persons from the upper classes [above \$4,000 annual family income] are more likely to receive psychiatric treatment . . . than persons who are either married or from the lower class" (Gove and Howell, 1974:98). The conclusion calls for an examination of *rates* for marital and income categories in which the nonhospitalized population is the denominator, but Gove and Howell give results for the relationship between marital-income status and symptom severity for patients who are hospitalized. Also, in their analysis, severity of impairment is the dependent variable; it is not a control variable. Consequently, their study does not show a relationship between resource variables and hospital admission with severity of illness controlled; this would require at least three variables, and they have only two. Moreover, by the nature of their study, it is not possible to address questions about the correlates of different types of societal reaction, since there are no data on type of societal reaction.¹

Introduction of the voluntary-involuntary variable solves these problems. Since the variable reflects two types of societal reactions, a relationship

¹ For criticisms of Gove and Howell's study, see Krohn and Akers (1977:348-9). This is not to say that Gove and Howell fail to show that resource variables and severity are related to hospitalized patients. They do show this, and their conclusion of a positive relationship between resource variables and hospitalization is plausible and may be true—even though we cannot be sure in the absence of population rates.

between status resources and type of admission would indicate that status resources are associated with type of societal reaction. Study of population rates is unnecessary because types of reaction that result in hospitalization are the focus of study; hence hospitalized patients are an appropriate population on which to conduct the study of relationships between status resources and societal reactions. Then after determining the nature of the relationship, and if data on severity of impairment are available, the relationship between status resources and societal reaction may be examined with degree of impairment controlled.²

Therefore, by investigating the relationship between socioeconomic status and type of hospitalization while controlling for psychiatric status, the results reported in this paper hopefully will contribute to a resolution of the debate about this particular aspect of the societal reaction or labeling perspective. As will be noted in the concluding section, however, data do not

bear on the *relative* validity of this perspective vs. the psychiatric perspective—support for one does not reject necessarily the validity of the other. Data are relevant for the validity of the societal reaction perspective only, and even then only so far as this perspective is relevant to the status resource hypothesis.

STATUS RESOURCE HYPOTHESIS

Although the status resource hypothesis has been tested in connection with mental hospital admissions, the hypothesis is more general. It stipulates that persons with more social and economic resources are better able to implement their life goals and achieve valued outcomes. Higher status persons usually have more income to purchase more and better goods and services that contribute to the style of living that the individual wishes to achieve, and that can be used to maintain that style in times of personal and social crises. High status persons also are apt to be more knowledgeable about the operation of the dominant community agencies and institutions and how the processes of these agencies and institutions can be used in one's behalf. They usually have greater access to individuals who are centrally connected in those agencies and institutions who can be called upon to help in the implementation of personal goals or to provide assistance in time of need. Thus, in general, individuals with higher socioeconomic status have greater resources to attain and maintain desired social states. This hypothesis is appropriate especially to the study of mental hospital admissions and in fact has given rise to two seemingly contradictory predictions about such admissions.

The first pertains to the distinction between voluntary and involuntary admissions. Since involuntary hospitalization has a number of undesirable consequences not associated with voluntary hospitalization, the hypothesis predicts that individuals with status characteristics which provide them with more rather than fewer resources are more apt to be hospitalized voluntarily, whereas individuals with fewer resources are more apt to be hospitalized involuntarily. Therefore, in-

² Rushing and Esco (1977) do find that involuntary male patients have fewer marital resources than voluntary patients when severity of impairment (mild, moderate, and severe) is controlled. Rice (1977) reports that a dichotomous measure of occupational status is correlated with type of admission when diagnosis (schizophrenia, alcoholism, neurosis, and sociopathy) is controlled. However, males and females are not distinguished in the analysis. It is possible, however, that there are substantial sexual differences in occupational status as well as type of admission. Consequently, the relationships reported may be due to sex being related either to the independent or dependent variables, or both. Gove and Fain (1977) report on the relationships of type of admission to degree of impairment and marital status, education, and income; their results are mixed but on balance indicate that involuntary patients have fewer resources and have more severe impairments. However, the relationship between resource variables and type of admission is not examined with severity of impairment controlled. Also, results are not reported separately for males and females. (In discussing the relationship of instrumental performance and type of admission, Gove and Fain do distinguish between males and females and find that for both sexes type of admission is associated with employment-unemployment status prior to admission, although the relationships are not significant statistically. Since the employed, by virtue of being employed, have more social and economic resources than the unemployed, these findings are consistent with the status-resource hypothesis. The analysis is limited, however, by the fact that only 9 of 169 female patients were employed.)

voluntary admissions relative to voluntary admissions are expected to increase as socioeconomic status decreases.³ This, of course, is the hypothesis to be investigated in this paper.

The second prediction, advanced by Gove and Howell (1974), is that social resources are significant in mental hospitalization because they facilitate the early entry of impaired individuals into hospitals. Persons of higher income are better able to obtain needed treatment quicker, and the behavior of persons with social resources (e.g., a wider network of social relations) may be more visible to others who help to get the individual into treatment, even if involuntarily. Thus a relationship is anticipated between socioeconomic status and length of time between onset of illness and when one enters treatment.

Both hypotheses may be true; verification of one need not falsify the other. Whereas the focus of Gove and Howell's hypothesis is on how quickly individuals get treatment, the focus of the other is on the way individuals are hospitalized. One hypothesis thus views promptness of treatment as the dependent variable; the other views type of process involved to incarcerate as the dependent variable. Patient socioeconomic status may be in-

³ This does not mean that voluntary admissions are viewed as controls for involuntary admissions. Gove and Fain (1977:17) believe

a careful analysis of the work by labelling theorists concerned with mental illness indicates that they generally are not concerned with hospitalization per se but only with involuntary hospitalization. In fact, the labelling theorists who do look at voluntary patients use them as controls against which to compare involuntary patients, and they argue that such a comparison provides support for labelling theory.

In the current version of the status resource hypothesis, as stated by Rushing (1971) and Rushing and Esco (1975), interest is in hospitalization per se because different types of hospitalization represent different types of societal reactions. One type is not viewed as a control for the other; rather each is viewed as an aspect of the same variable, degree of coercion in societal reactions. The focus is on the relationship between socioeconomic status resources and that variable. (Gove and Fain's reference is apparently to Scheff [1974:499] who suggests that the ratio of involuntary to voluntary admissions may control for degree of illness.)

versely related to promptness of treatment when type of admission is controlled (that is, socioeconomic status and promptness of treatment may be related for voluntary and involuntary admissions); and socioeconomic status may be related to type of admission when duration of symptomatology is controlled. Since we do not have data on the duration of symptomatology, our concern is exclusively with the effect of patient psychiatric status on the possible relationship between socioeconomic status and type of hospital admission. Our hypothesis is that socioeconomic status has an effect even after psychiatric status is controlled. In addition to investigating the general status resource hypothesis (the main effects hypothesis), two hypotheses which stipulate that socioeconomic status interacts with psychiatric status will be investigated.

Interaction of Socioeconomic Status and Psychiatric Diagnosis

One interaction hypothesis states that the relationship between status resources and type of hospital admission is stronger for functional than for organic disorders. Although methodological problems of psychiatric diagnosis are well-known, Torrey (1974:67) states that most organic disorders are rather easily identified, even by laymen with little training. In addition, organic disorders are often based on laboratory tests and medical examinations which yield very objective (that is, reliable) data (e.g., x-rays which reveal structural brain damage). Nonorganic disorders, however, are inferred altogether from the individual's behavior and the reports of others, which are probably subject to more variable interpretation than data based on tests for disorders with a physical base. In short, diagnoses of organic disorders may be less ambiguous than diagnoses of nonorganic disorders. Consequently, nonpsychiatric and non-medical factors, such as status resources, may play a larger role in the societal reaction and how an individual is hospitalized for mental illness. Hence a stronger relationship would be expected between socioeconomic status and type of admission for functional disorders.

Interaction of Socioeconomic Status and Impairment

Traditional sociological conception (as opposed to the newer societal reaction conception) of deviant behavior emphasizes the deviant character of the individual's behavior, in contrast to the nature of the societal reaction. The significance of behavior for the societal reaction should not be dismissed in the societal reaction perspective, however, since it is probable that the nature of the reaction depends to some extent on the nature of the individual's behavior (Rushing and Esco, 1977). Specifically, the more disruptive and disorganized the individual's behavior to the community (the more aggressive, manic and bizarre the behavior), the more severe and forceful the reaction is apt to be. At the extreme, individual status may have no effect on the societal reaction; extremely disruptive persons are apt to be dealt with coercively *regardless* of their status characteristics. Status characteristics are more important as contingencies when the individual's behavior is less disturbing (Rushing and Esco, 1977).

Although a direct measure of behavior is not available, a measure of severity of patient condition upon admission, in terms of degree of impairment (none, minimum, mild, moderate, severe) is available. This is an indirect measure of degree of behavioral deviance because assessment of the seriousness of an individual's psychiatric condition is inferred from statements and behavior by the individual and by statements of others concerning the individual's verbal and non-verbal behavior. Consequently, while this may be a valid assessment of the severity of the individual's mental disorganization and personal discomfort, it is also an index of degree of deviance—the level of disruption, trouble, and concern an individual causes others. Thus we would expect the relationship between educational status and type of hospital admission to be stronger for the minimally impaired than for those who are severely impaired.

The fact that assessments were made after admission may have been influenced by type of admission since many psychi-

atrists apparently believe that involuntary patients are sicker (Gove and Howell, 1974) and, as noted, some evidence supports this belief (Rushing and Esco, 1975; Gove and Fain, 1977). Nevertheless, assessed impairment probably is influenced to a substantial degree by professional psychiatric judgment and is not completely based on whether the individual was committed by the court, as an extreme labelling position might suggest. Even so, if assessed impairment is influenced by type of admission, its use as a control variable for the relationship between socioeconomic status and type of admission actually biases findings against the status resource (main effects) hypothesis. This is so because control for deviance-impairment would also control for the dependent variable. This would reduce the relationship between socioeconomic status and type of hospital admission. Therefore, test for the status resource hypothesis is conservative.

DATA

The data are for all 20–64 year-old first admissions to each of six mental hospitals in Tennessee who were admitted and discharged during the ten-year period between fiscal 1956 and 1965. There were a total of 11,052 patients. Upon admission, a variety of information was collected on patients. This included education, diagnosis, age, and sex, as well as level of impairment and type of admission. Education will be used to index socioeconomic status and is categorized by number of years of education as follows: 0–4, 5–7, 8, 9–11, 12, and above 12. Diagnosis is in terms of the traditional psychiatric categories as prepared by the Committee on Nomenclature and Statistics of the American Psychiatric Association (American Psychiatric Association, 1956). Disorders diagnosed as organic or toxic in origin are distinguished from nonorganic or functional disorders in order to test the first interaction hypothesis. Functional disorders include psychosis, neurosis, and personality disorders. As for level of impairment, the first three categories on the impairment scale (none, minimal, and mild) are combined into one category,

mild. This variable will provide the basis for testing the second interaction hypothesis. Test for the main effects hypothesis will include psychiatric diagnosis and level of impairment as control variables.

Although there are a total of 11,052 patients (5,283 males and 5,769 females), missing or inappropriate data on some variables preclude analysis from being based on all patients. Patients diagnosed as retarded (380) were eliminated since retardation is not a mental illness diagnosis; 239 were not given a diagnosis and 170 were given a diagnosis other than those designated above. Finally, 2,895 (26.2%) were not given an assessed impairment rating. Why so many patients were not given an impairment rating is not known, but in light of the very small percentage who were not given a diagnosis (2.2%), it could be due to doubt on the part of hospital staff as to the seriousness of the disorder.

ANALYSIS AND FINDINGS

Before examining the data as they pertain specifically to the hypotheses, several observations are in order. First, the two versions of the status resource hypothesis imply different conceptions about the role of state mental hospitals—as treatment centers or as institutions of social control. If the former were dominant, most patients would enter hospitals in order to receive help and treatment, and presumably they would enter voluntarily, as is the case for most nonpsychiatric hospitals. However, if the hospitals were primarily places for custodial care, safekeeping and the protection of society, we would expect most to enter reluctantly; at the extreme, hospitals would approximate prisons, in which case virtually no one would enter voluntarily.

Results show that by this criterion, state mental hospitals in Tennessee would not be considered exactly as prisons, since 28.2% entered voluntarily. At the same time, 72.8% did not enter of their own will but had to be legally coerced into doing so. This does not include prison transfers and criminal court commitments. (Males were slightly more apt to be committed, 74.8% vs. 68.9%). It would seem, then,

that the custodial-social control function in which the protection of society is central was a major function of these hospitals.⁴

Table 1 shows the relationship of education to type of admission in percentage for males and females who are involuntarily committed. It is clear that the lower status patients were more apt to have been committed by the courts. The effects of age need to be examined, however, since age of patient may be related to education and to type of admission. Analysis reveals

⁴ This appears to be the case for hospitals in other states during the period. In California for 1958, 81.8% of all voluntary admissions and civil commitments were by commitment (Morgan and Cook, 1963). For Washington state, Rushing (1971) reports that 60.2% of all first admissions for 1955–64 were committed and Linsky (1971) reports 70.9% for the period 1957–63. (Apparently the difference reported by Linsky and Rushing stems from the fact that Rushing includes only patients who were voluntary admissions and commitments for mental illness, thus excluding sexual psychopaths and psychopathic delinquents who were also committed by the courts as well as other cases involving court intervention, such as sex deviants, delinquent behavior, and other illegal activities.) In North Carolina for 1970–73, Rice (1977) reports that 50.0% were committed by the courts. Thus figures for entire states during 1955–64 and 1970–73 show that a large proportion of patients were committed by the courts. This suggests that state mental hospitals in Tennessee were not atypical; individuals in other states were apparently reluctant to enter state hospitals. A much lower percentage (33.3%) of patients were committed in Gove's study (Gove and Fain, 1977) for one Washington state hospital for the period June, 1962–December, 1964. This may be the result of Gove's patients coming from only one county. The small number of patients (258) may not have been typical of the patient population. These patients were admitted to a demonstration project in one state hospital, and an unusually high percentage (66.7%) were females (percentages of females in other studies are 53.4% [Morgan and Cook, 1963] and 49.5% [Linsky, 1971]; in the current study 52.9% are female). All studies report higher proportions of involuntary patients for females, but the percentage of voluntary patients is much higher for both males and females in Gove's study. In comparison with 53.5% (males) and 73.8% (females) in Gove's study (Gove and Fain, 1977), Morgan and Cook (1963) report 15.1% (males) and 21.0% (females), Linsky (1971) reports 21.9% (males) and 36.3% (females); figures for the current study are 25.9% (males) and 31.4% (females). These sexual differences are consistent with the findings by Horowitz (1977) that females are more apt to seek help for psychiatric problems than are males. Sex differences in type of admission are dealt with in detail in a forthcoming paper.

Table 1. Relationship between Socioeconomic Status (Occupation and Education) and Type Admission by Sex, in Percentage of Court Commitments (21-64 Year Old First Admissions to Tennessee State Hospitals, 1956-65)

	Education (Years)						
	0-4	5-7	8	9-11	12	Above 12	
Males	87.9 (794)	82.8 (998)	78.9 (920)	69.8 (827)	58.9 (725)	45.6 (407)	-.41*
Females	86.7 (473)	81.1 (994)	78.1 (1,105)	59.1 (1,129)	57.1 (1,097)	50.9 (383)	-.37*

Note: Figure in parenthesis is N.

* $P < .000001$.

that the gammas between age and type of admission are only $-.11$ for males and $-.23$ for females. (In this analysis age is categorized into five-year intervals, 20-24 . . . 60-64.) The partial gammas between type of admission and education are $-.41$ (males) and $-.35$ (females); comparisons with gammas in Table 1 indicate that the relationship of status resources to type of hospitalization is not a function of age.

Diagnosis is correlated with behavioral deviance-impairment; gammas are $.33$ for males and $.30$ for females, with organic disorders being more apt to be assessed as severely impaired. Both of the variables are correlated with educational status and type of admission. For diagnosis and education, gammas are $-.30$ and $-.31$ for males and females, respectively, while for type admission they are $.33$ and $.45$ (with organic disorders being more apt to be involuntarily admitted). Findings are similar for deviance-impairment: gammas are $-.15$ and $-.14$ (diagnosis); and $.43$ and $.55$ (type admission). It remains to be seen, however, whether psychiatric status and deviance-impairment account for the relationship between socioeconomic status and type of admission and thus eliminate the main effect of education, and, also, whether educational status interacts with psychiatric status and degree of deviance as our hypotheses stipulate.

The linear probability estimation model developed by Grizzle et al. (1969) will be used in testing these hypotheses. This model is similar to regression analysis using dummy variables in that it makes a prediction for the dependent variable for each value of the independent variable. It differs in that values can be predicted for a

dichotomous dependent variable. The technique will thus allow for the estimation of the effect of education—controlling for psychiatric diagnosis and behavioral deviance—and further allows us to test whether this direct effect of education differs in magnitude depending on diagnosis and behavioral deviance as suggested by the two interaction hypotheses.

In using the technique, effects coding rather than dummy coding was used.⁵ As in dummy coding the number of coefficients for each variable in the equation is $k - 1$; for example, in the present instance, there would be two coefficients for impairment since there are three categories (mild, moderate and severe). Unlike dummy coding, however, in effects coding the sum of the codes (and hence the sum of the coefficients) for a particular variable equals zero. Thus, in effects coding, the value of the coefficient for the missing category for each variable in which $k - 1 \geq 2$ equals the value which sets the sum of all the coefficients to zero. The general rule is: $k - 1$ coefficients plus the unstated coefficient is equal to zero. (In dummy

⁵ In effects coding, a dichotomous variable such as diagnosis is scored as 1 (functional) and -1 (organic) rather than 0 and 1 for each diagnosis depending on the presence or absence of one or the other diagnosis, as in dummy coding. For variables with more than two categories, dummy coding scores all categories except one ($k - 1$) as 0 or 1 depending on the presence of the characteristic. For example, mild and moderate would each be scored as 0 or 1 depending on whether the individual were assessed as mildly or moderately impaired, such that a mildly impaired individual would be scored as 1 (for mild) and 0 (for moderate). In effects coding, scoring for impairment is 0 or 1 (mild), 1 or 0 (moderate), and -1 and -1 (severe).

coding it is the missing coefficient that is equal to 0.0, so it is the excluded category that is the reference category rather than the sum of the codes as in effects coding.) The three variables (education, diagnosis, and deviance-impairment) and the two interaction terms yield the following equation:

$$\begin{aligned}
 IV = & C + B_1E_1 + B_2E_2 + B_3E_3 \\
 & + B_4E_4 + B_5E_5 + B_6D_1 + B_7I_1 \\
 & + B_8I_2 + B_9(E_1D_1) + B_{10}(E_2D_1) \\
 & + B_{11}(E_3D_1) + B_{12}(E_4D_1) + B_{13}(E_5D_1) + B_{14}(E_1I_1) \\
 & + B_{15}(E_2I_1) + B_{16}(E_3I_1) \\
 & + B_{17}(E_4I_1) + B_{18}(E_5I_1) \\
 & + B_{19}(E_1I_2) + B_{20}(E_2I_2) \\
 & + B_{21}(E_3I_2) + B_{22}(E_4I_2) \\
 & + B_{23}(E_5I_2) + e,
 \end{aligned} \quad (1)$$

where

IV = proportion of patients involuntarily admitted;

C = constant;

E = educational status ($E_1 = 0-4$ years, . . . $E_5 = 12$ years; > 12 years omitted);

D = diagnosis ($D_1 =$ functional; organic omitted);

I = impairment ($I_1 =$ mild; $I_2 =$ moderate; severe omitted);

e = error.

The number of coefficients can be reduced if the differences between coefficients for the categories of a variable with more than two categories are equal. For example, if the differences between coefficients for education (B_1E_1 . . . B_5E_5) were approximately equal, education may be considered a continuous variable and the difference between categories treated as equal intervals. Analysis shows that this is the case for education since all differences between categories are approximately .06 for males and .05 for females which means that with all other terms controlled, on the average, the percentage of involuntary admissions increases 6% and 5% with each internal decrease in education. Since chi-square test reveals that the difference between the predicted value and actual value for education is not significant statistically at the .05 level for males or

females, education may be treated as a continuous-variable with equal intervals rather than as a categoric variable and is scored as $-2.5, -1.5, -0.5, 0.5, 1.5$ and 2.5 . Equation (1) may therefore be rewritten as follows:

$$\begin{aligned}
 IV = & C + B_1E + B_2D_1 + B_3I_1 + B_4I_2 + \\
 & B_5(ED_1) + B_6(EI_1) + B_7(EI_2) + e.
 \end{aligned} \quad (2)$$

Findings are presented in Tables 2, 3, and 4.

Table 2 gives the predicted and actual proportion of involuntary admissions for each education-diagnostic-impairment category by sex. In most instances the correspondence between the predicted and actual value is quite close. Although the correspondence is not perfect and in certain cells the difference is substantial, such differences are found primarily in those cells containing a relatively small number of cases.

Our theoretical model stipulates that societal reactions are significantly influenced by three variables: socioeconomic resources (educational status), interaction of socioeconomic resources and diagnosis, and interaction of socioeconomic resources and degree of deviance-impairment. These terms along with diagnosis and impairment are included in equation (2). Interaction between diagnosis and impairment is excluded (as are all higher-order interactions) since there is no theoretical reason why the effect of diagnosis should vary depending on severity, and in any case, it is of no particular relevance to the status resource hypothesis and the two corollaries.⁶

If the model in equation (2) accounted for the data perfectly, there would be no statistically significant differences between the actual and predicted values (there would be no error) in Table 2. The linear probability estimation model permits a test of such differences. Analysis shows that $\chi^2 = 43$ for males ($p < .03$; 28

⁶ The inclusion of the interaction between D and I would nevertheless impose a more stringent condition for testing the hypotheses, since control for (DI) might cause some of the coefficients for theoretically significant relationships to reduce to 0.0. This term is included in the analysis and is reported on below.

df.) and 47 for females ($p < .02$; 28 df.). Thus, our model is not a perfect fit for the data. Comparisons of the observed and predicted proportions in Table 2 demonstrate, however, that the model accounts for the data to a large degree, especially for our substantive purposes. This is especially clear from the following.

In the linear-probability estimation model, chi-square test statistics are computed for each variable and partitioned into components analogous to the analysis of variance. This allows for the simultaneous test for the effects of education, diagnosis, and deviance-impairment and the two interaction terms each with the other variables controlled. Results are reported in Table 3. The table shows that the error chi-square of 43 and 47 would increase by 156 (males) and 87 (females) if education were eliminated from the analysis, and 19.8 and 19.4 (males) and 25.4 and 8.9 (females) if the two interaction terms were excluded. Thus, the main effects hypothesis receives support even when diagnosis, deviance-impairment, and the interaction terms are controlled, and each of the two interaction hypotheses receive support after all three basic variables and the interaction term in the other interaction hypothesis in the model are controlled. Further, all three relationships remain significant even after the other first-order interaction term, diagnosis and deviance-impairment, is included in the analysis and hence con-

Table 3. Chi-square Values for the Relationship of Societal Reaction (Type of Hospital Admission) and Five Variables (All 21-64 Year Old First Admissions to Tennessee Mental Hospitals, 1956-65)

Variable	Males	Females	Degrees of Freedom
Status resources-education (E)	156*	87*	1
Diagnosis (D)	45*	72*	1
Deviance-impairment (I)	198*	386*	2
ED	19.8*	25.4*	1
EI	19.4*	8.9*	2

* $P \leq .01$.

trolled (results not shown). (This interaction term is not significant statistically.) In general, then, results support the status resource hypothesis and the two hypotheses concerning the specification of the effects of status resources. Additional analysis allows the specification of these effects still further.

First, Table 2 shows the average effect on type of admission per interval change in education separately for twelve categories of diagnosis, impairment, and sex. These are analogous to partial regression coefficients; in this case the effect of education with diagnosis and impairment are controlled. This can be seen from Table 4 which presents the coefficients for each of the terms in equation (2), plus the two coefficients that are unstated. As

Table 4. Coefficients Based on the Linear Probability Model for Each of Eight Terms Plus the Unstated Terms in Equation (1) (21-64 Year Old First Admissions to Tennessee Mental Hospitals, 1956-65)

	Coefficients		Standard Error	
	Males	Females	Males	Females
Constant	.74	.73	.01	.01
Education (B_1)	.06	.05	<.01	.01
Diagnosis (B_2)	-.05	-.07	.01	.01
Impairment				
Mild (B_3)	-.10	-.19	.01	.01
Moderate (B_4)	-.04	-.03	.01	.01
Severe	.14	.16	.01	.01
Education X Diagnosis (B_5)	.02	.02	<.01	<.01
Education X Impairment				
Mild (B_6)	.01	.02	.01	.01
Moderate (B_7)	.01	.00	.01	.01
Severe	-.02	-.02	.01	.01

theorized, education has an effect on the probability that one will be involuntarily rather than voluntarily hospitalized, even when degree of impairment and type of diagnosis are controlled. With the control variables held constant, on the average, as the educational level of men increases by a category, the probability of involuntary commitment drops by .06 (.05 for women). But as predicted, the magnitude of the direct effect of education differs depending on diagnosis and assessed impairment. Thus, as Table 2 clearly shows, for males and females, direct effect of education is small (.012) for cases with severe impairment and organic disorders, and large (.09) where mild impairment and functional disorders are in question.⁷

Note from Table 2 that the average effect of education is about the same for the mild and moderate categories for functional and organic diagnoses among males and differs only for the severe category. This suggests that in the case of behavioral deviance, at least among males, the effect of status resources are modified by behavior only when behavioral deviance is extreme. Table 4 shows that only in the case of the severe category is the interaction between impairment and education statistically significant, since only in this case is the coefficient at least twice its standard error. This is true only for males, however, since for females the coefficients for both the mild and severe categories are significant statistically. Still, for both males and females, the coef-

ficient for the severe category is significantly different from the other two coefficients: for males the coefficients of .01 for the mild and moderate categories are beyond the range (at the .01 level) of $-.02 \pm .01$, and among females the coefficients of .02 and .00 fall beyond the range of $-.02 \pm .01$.

Thus, to conclude, for males and females, when behavior is extremely deviant (very disruptive and bizarre), behavior modifies the effects of status resources on the societal reaction. The effect may not be limited to the interaction of education with the extremely deviant vs. the less than extremely deviant among females; the effect is significant at both ends of the impairment scale. Results in Table 2 indicate that the effects of education on societal reactions decrease continuously from the mild to the severe category for functional and organic diagnoses alike. Results for males are similar to those reported by Rushing and Esco (1977) for marital status. The effects of social resources (with the married having the most resources and single having the least) are modified only among persons who are extremely deviant. The precise nature of the interaction of behavioral deviance and status resources on the societal reaction may vary between males and females. In any case, for both males and females, there is an overall interaction effect as one of the interaction hypotheses stipulates.

Finally, although it is the distinction between organic and functional diagnoses that is particularly relevant to the present paper, diagnostic distinctions within the functional category (psychoses, neuroses, and personality disorder) also may be significant. Specifically, the more severe the diagnosis (psychosis being most severe, personality disorder least severe), the more coercive the societal reaction may be. Therefore, analysis identical to that in Tables 2, 3, and 4 was conducted for psychoses, neuroses, and personality disorders. Results show that psychotics were most apt to be involuntarily hospitalized and persons with personality disorders least; coefficients are .16, $-.16$, and .00 for males and .18, $-.10$, and $-.08$ for females. Our interest, however, is not in

⁷ Observe that from Table 4 the coefficient for the interaction of education and diagnosis for males is .02; this means that .02 is to be added to the coefficient for education (.06), for a value of .08. Observe further that the coefficients for interaction between education and impairment are .01 (mild), .01 (moderate), and $-.02$ (severe). Hence, for the functional-mild category, the effect would be $.06 + .02 + .01 = .09$, which is what Table 2 shows. The value for the functional-moderate category is also .09. For the functional-severe category, the value would be $.06 + .01 + (-.02) = .05$, which is the average predicted value in Table 2. Among organic disorders, for the mild category, from Table 4 we have $.06 + (-.02) + .01 = .05$; the same value obtains for the moderate category. For the severe category, the value would be $.06 + (-.02) + (-.02) = .02$. With the exception of the differences due to rounding errors, the correspondence between Table 2 and Table 4 is exact. Similar computations may be made for females.

whether these distinctions are related to type of hospitalization, but in whether they eliminate or modify the effects of education. Analysis shows that they do not. For males and females, coefficients for education are significant after diagnosis is controlled, and there is no significant interaction between education and functional diagnoses. It would seem, therefore, that diagnostic distinctions within the functional category are of little or no relevance insofar as the status resource hypothesis is concerned. At the same time, however, *if* these distinctions are reliable and valid, they would seem to call into question the hypothesis which stipulates an interaction effect between status resources and deviance-impairment. This aside, results consistently support the view that status resources are significant contingencies in the societal reaction.

CONCLUSION

Within the limits of the data—they are for only one state and one period of time—results help to resolve the controversy about the influence of social attributes and, especially, of status resources in the mental hospitalization process, since they rather conclusively show that status resources do have such an effect. At the same time, results neither deny the role of psychiatric status in the hospitalization process. (As we have seen, diagnostic distinctions are related to type of hospital admission.) Nor do the results deny that most or even that some patients are mentally ill rather than just labelled as such. In this paper we have not tested the relative validity of the societal reaction perspective and the psychiatric perspective which states, as I understand it, that patients enter mental hospitals because they are sick (Gove and Howell, 1974:86). To affirm the status resource hypothesis does not in itself reject this statement. Moreover, to contend that individuals are hospitalized because they are ill and engage in deviant behavior *or* because they have certain social characteristics is to oversimplify. As we have seen, illness-deviance and social characteristics interact.

Finally, the status resource hypothesis and the data presented in support of that hypothesis are relevant to only one aspect of the societal reaction framework. One critic of the perspective claims that the two central issues in the perspective are (1) the nature of the behavior that led to the societal reaction and (2) the subsequent effect of the societal reaction in the careers of deviants (former mental patients) (Gove, 1976a; Gove and Fain, 1977:676). These are indeed important *general* issues,⁸ but they do not exhaust all important issues within the perspective (see, for example, Conover, 1976). Certainly we have dealt with neither of them in this paper. We have not dealt with the causes of individuals being labelled mentally ill in the first place—psychiatric or socioeconomic status. Instead, we have been concerned with accounting for different types of societal reactions to individuals who are subsequently hospitalized; all are labelled to some extent simply by virtue of being hospitalized. And all (or none) may be truly ill. In addition, in no instance do we address the question of the causes of deviant careers, whether this is labelling or something else. We have dealt only with social contingencies of the societal reaction, and, even

⁸ I do not agree with the rather narrow wording, namely, that "labelling theory holds that the prime cause of being labelled a deviant is the person's marginal social and economic attributes and not the degree and severity of the person's deviant behavior" (Gove and Fain, 1977:676). The statement is not inclusive enough. To marginal social and economic attributes should be added group tolerance, visibility of behavior, disruptiveness of behavior (Scheff, 1966; Mechanic, 1968:438-9) as well as the economic costs of treatment and the availability of treatment and institutional facilities (which would include level of crowding in institutions). Public temper and cultural climate are also important. (We know, for example, that involuntary hospitalization is less acceptable to psychiatrists, legislators and the public today than in years past, and the percentage of patients committed by the courts has decreased accordingly.) Also, even a more inclusive list of non-deviant attributes would not make the above quote a correct interpretation of the position of some labelling theorists. For example, while Lemert (1951) emphasizes the societal reaction and its effects in the generation of secondary deviance, he does not hold that the degree and severity of the person's deviant behavior is unimportant in eliciting the societal reaction (see Lemert, 1976:244-5).

here with only one particular contingency—status resources. Other possible contingencies (Scheff, 1966), including contingencies that may be even more important than status resources, such as visibility of residual deviance, group and community tolerance for residual deviance, and economic costs of treatment have not been investigated. And we have not examined all types of possible reactions to persons who are residually deviant. We have only addressed the question of whether coercive sociolegal reactions in mental hospitalization are associated with the status resources of individuals who are considered to be residual deviants. Evidence consistently shows that they are.

REFERENCES

- American Psychiatric Association
1952 Nomenclature and Statistics of the American Psychiatric Association. Washington, D.C.: American Psychiatric Association Mental Hospital Service.
- Conover, Patrick W.
1976 "A reassessment of labelling theory: a constructive response to criticism." Pp. 228–43 in Lewis A. Coser and Otto N. Larsen (eds.), *The Uses of Controversy in Sociology*. New York: Free Press.
- Ennis, Bruce and Loren Siegel
1973 *The Rights of Mental Patients*. New York: Avon.
- Gove, Walter R.
1975 "The labelling theory of mental illness: a reply to Scheff." *American Sociological Review* 40:242–8.
1976a "Deviant behavior, social intervention, and labelling theory." Pp. 219–27 in Lewis A. Coser and Otto N. Larsen (eds.), *The Uses of Controversy in Sociology*. New York: Free Press.
1976b "Reply to Imershein and Simons (1976) and Scheff (1975)." *American Sociological Review* 41:564–7.
- Gove, Walter R. and Patrick Howell
1974 "Individual resources and mental hospitalization: a comparison and evaluation of the societal reaction and psychiatric perspectives." *American Sociological Review* 39:86–100.
- Gove, Walter R. and Terry Fain
1977 "A comparison of voluntary and committed psychiatric patients." *Archives of General Psychiatry* 34:669–76.
- Grizzle, James E., C. Frank Starmer and Gary G. Koch
1969 "Analysis of categorical data by linear models." *Biometrics* 28:137–56.
- Horowitz, Allan
1977 "The pathways into psychiatric treatment: some differences between men and women." *Journal of Health and Social Behavior* 18:169–78.
- Krohn, Marvin B. and Ronald L. Akers
1977 "An alternative view of the labelling versus psychiatric perspective on societal reaction to mental illness." *Social Forces* 56:341–61.
- Lemert, Edwin M.
1951 *Social Pathology*. New York: McGraw-Hill.
1976 "Response to critics: feedback and choice." Pp. 244–9 in Lewis A. Coser and Otto N. Larsen (eds.), *The Uses of Controversy in Sociology*. New York: Free Press.
- Linsky, Arnold S.
1970 "Who shall be excluded: the influence of personal attributes in community reaction to the mentally ill." *Social Psychiatry* 5:166–71.
- Mechanic, David
1968 *Medical Sociology*. New York: Free Press.
- Morgan, Richard D. and Leonard R. Cook
1963 "Relationships of methods of admission to length of stay in state hospitals." *Public Health Reports* 78:619–29.
- Rice, Willy E.
1977 "Mental hospitalization and individual resources: a consideration of socio-legal factors in evaluating the efficacy of the psychiatric and societal reaction perspectives." Unpublished manuscript.
- Rushing, William A.
1971 "Individual resources, societal reaction and hospital commitment." *American Journal of Sociology* 77:511–26.
- Rushing, William A. and Jack Esco
1975 "The status resource hypothesis and length of hospitalization." Pp. 445–55 in William A. Rushing (ed.), *Deviant Behavior and Social Process*. Rev. ed. Chicago: Rand McNally.
1977 "Status resources and behavioral deviance as contingencies of societal reaction." *Social Forces* 56:132–47.
- Scheff, Thomas J.
1966 *On Being Mentally Ill: A Sociological Theory*. Chicago: Aldine.
1974 "The labelling theory of mental illness." *American Sociological Review* 39:444–52.
- Torrey, E. Fuller
1974 *The Death of Psychiatry*. Radnor: Chilton.

IS STATUS ATTAINMENT RESEARCH ATHEORETICAL?*

PATRICK M. HORAN

University of Georgia

American Sociological Review 1978, Vol. 43 (August):534-541

Recent critiques of status attainment research have emphasized the atheoretical character of that research. This paper develops an opposing view, based on the premise that status attainment research is theory-laden. The distinctive characteristics of status attainment research which have provoked outside criticism are shown to be derivative of a neoclassical, functionalist conception of social structure.

Two recent critiques (Coser, 1975; Burawoy, 1977) of status attainment have based their criticisms of the direction and substantive content of status attainment research on the premise that this research tradition is atheoretical, or as Coser (1975:691) puts it, a "method in search of substance." I wish to suggest that in treating status attainment as an atheoretical product of mindless methodology, these critics have reduced the effectiveness of their critique of status attainment. By failing to identify the theoretical model underlying status attainment research, they have put undue emphasis on the methodological differences between status attainment and other research traditions in stratification and have ignored important theoretical differences.

Status attainment is not atheoretical. Quite the contrary, it is heavily theory-laden (Hanson, 1972). And it is this theory-laden character which lies at the root both of the notable successes of status attainment research at solving problems which it defines, and the striking failures of status attainment research to address problems which are substantively interesting to proponents of other theoretical perspectives.

Before attempting to use Hanson's concept of "theory-laden" to interpret the status attainment research tradition, it should be noted that Hanson's view of the role of observations in empirical research

differs from the positivist conception which dominates the sociological literature. In the latter, theory is represented as a set of propositions which may be tested against a body of theoretically neutral empirical observations or facts. In contrast Hanson maintains that systematic empirical observations are conditional on an interpretational framework which attaches meaning to, and thus legitimates, these observations as "facts." This unavoidable use of theory to organize and direct our empirical inquiry leads to a situation in which research based on such observations can be seen as theory-laden. Gordon (1972:27) has taken a similar position. He suggests that interpretation of an empirical literature be approached by a procedure of "revealed theoretical preference" in which "manifest 'bundles' of empirical analysis help 'reveal' the application of a set of theoretical assumptions. . . ."

THEORY AND OCCUPATIONS

Most conceptions of industrial society give occupation a major role in the social organization of society. Let us presume a primitive definition of occupation: a distinct social position defined in terms of characteristic activities in the socioeconomic realm. This primitive representation will play the same role in my argument as the lines of a Gestalt drawing play in Hanson's. That is, it provides a "neutral" benchmark against which distinct theoretical observations or measurements of occupation may be seen. Such neutral observational schemes are, of course, purely fictional, but have some utility in

* I would like to thank Margaret Farnworth, Mary K. Killian and Charles M. Tolbert II as well as the anonymous reviewers for their helpful comments and suggestions.

introducing Hanson's concept of theory-laden observations to the analysis of stratification research.

There is another, historical, justification for using a categorical occupational scheme as a point of reference in the discussion here. The mobility table research tradition which preceded status attainment employed categorical occupational schemes, and thus much of the explicit discussion of the representation of occupation in status attainment research involves comparisons with such categorical schemes. Given a primitive, categorical occupation classification, the introduction of theoretical considerations must serve to organize our observation and classification (i.e., measurement) of occupations according to theoretical criteria. For example, a Marxist application might represent occupational differences in terms of categories reflecting relations to the means of production and then use this measurement of occupation in analysis of the relationship between occupational position and selected individual behaviors (e.g., Wright and Perrone, 1977). Likewise, in status attainment research the measurement of occupations provides a key to the underlying theoretical model.

While the measurement operation provides an important key to understanding the theoretical foundations of a research program, we need not expect any widespread consciousness among researchers of explicit links between measurement operations and theoretical foundations. A researcher in an ongoing research tradition typically will not be called upon to think through the fundamental measurement operations and to justify them in theoretical terms. Indeed, Kuhn (1970:193) argues that just the opposite is the case:

One of the fundamental techniques by which members of a group . . . learn to see the same things when confronted with the same stimuli is by being shown examples of situations that their predecessors in the group have already learned to see as like each other and as different from other sorts of situations.

Thus among the practitioners in a research tradition, measurement typically is seen as the problem of extracting the "right"

piece of reality; less often is it seen as the imposition of a theoretical perspective.

The present inquiry into the theory-laden character of status attainment research will require answers to the following sorts of questions: Does status attainment research rest on an empirical base provided by a distinctive measurement operation? If so, what is the nature of that representation, and what basic theoretical perspective is required to justify such a representation? In general *what would our basic conception of stratification have to be in order for the status attainment model to be appropriate?*

STATUS ATTAINMENT AND OCCUPATIONS

What is the measurement operation which differentiates status attainment from the earlier mobility table research? It is the replacement of discrete occupational categories with numerical scores derived from the aggregate evaluations of those occupations by a set of sample respondents (Reiss, 1961). These evaluations have been projected from the original limited NORC sample of occupations to the census detailed occupational classification by Duncan (1961) and have been revised and updated by Siegel (1971). Thus in place of discrete occupational positions as measures of occupation, there are numerical levels of evaluation used to represent the positions of occupations in the social order. In this way our basic representation of occupations in the social order is changed from that of a set of discrete social positions to that of a continuum of presumably consensual popular evaluation in which differences between occupations can only be matters of degree. This implies that the important differences between occupations are differences in superiority/inferiority, and that such differences between occupations can be readily assessed by any and all participants in the society.

What an interesting conception of social structure, characterized by a popular consensus on the evaluation of occupations, by a unidimensional system of evaluation, and by numerical measures of occupational superiority/inferiority. On the one hand, how strange it is to have social



structure reduced to shared values among individuals. On the other hand, there is a certain familiar ring to this portrayal. In his "An Analytical Approach to the Theory of Social Stratification," Talcott Parsons (1940:841-4) writes:

Social stratification is regarded here as the differential ranking of the human individuals who compose a given social system and their treatment as superior and inferior relative to one another in certain socially important respects . . . It is only in so far as differences are treated as involving or related to particular kinds of social superiority and inferiority that they are relevant to the theory of social stratification. . . . Consideration of certain aspects of social systems described in terms of the theory of action shows readily why stratification is a fundamental phenomenon. In the first place, moral evaluation is a crucial aspect of action in social systems. . . . The second crucial fact is the importance of the human individual as a unit of concrete social systems. If both human individuals as units and moral evaluation are essential to social systems it follows that these individuals will be evaluated as units and not merely with respect to their particular qualities, acts, etc. . . . Unless there is to be a functionally impossible state of lack of integration of the social system, the evaluations by A and B of their associate C must come somewhere near agreeing; and their relative ranking of C and D must broadly agree where the necessity for comparison arises. . . . There is in any given social system, an actual system of ranking in terms of moral evaluation. But this implies in some sense an integrated set of standards according to which the evaluations are, or are supposed to be made. . . . The actual system of effective superiority and inferiority relationships . . . will hence be called the system of social stratification.

This statement of the functionalist conception of social stratification, a conception on which the more familiar Davis and Moore (1945) presentation rests, indicates with remarkable clarity the correspondence between the functionalist theoretical framework and the status attainment representation of occupational structure. The status attainment measurement model represents occupational differentiation as functionalism tells us it must be. This is not to say that this empirical measurement operation validates the

functionalist conception of a unidimensional, consensual evaluation of occupations. Reiss (1961) and Goldthorpe and Hope (1974) have demonstrated both the multiplicity of evaluation criteria which respondents use in rating occupations and the multidimensional character of the ratings obtained from respondents. Thus the occupational prestige scales on which status attainment rests are unidimensional *by construction*.

Few would deny that the measurement of occupation in terms of prestige (or status) constitutes a major change in the representation of occupational differentiation. It is one thing to note that prestige (or status) is one of several important aspects of occupational differentiation underlying, say, occupational mobility patterns between fathers and sons. It is quite another thing to say that henceforth, in the study of mobility, occupations will be represented as prestige units and nothing else. In presenting the basic regression application which lies at the heart of the methodology of status attainment, Duncan and Hodge (1963:630) argue that this application:

requires one fundamental assumption that is not intrinsic to several of the other techniques commonly used in studying occupational mobility. It must be assumed that occupations can be assigned values on the scale of a single quantitative variable. In point of fact, occupations differ in a number of ways, and the great variety of differences among kinds of work can by no means be reduced to variation along a single dimension.

In their presentation of the basic model of status attainment, Blau and Duncan (1967:117) provide a more positive view of the measurement model.

If the focus on vertical mobility, therefore, involves a simplification of the actual process by which individuals find their way into occupational roles, it is nonetheless a justifiable simplification. To study one aspect of a complex phenomenon is not to deny that other aspects exist.

But the effect of the growth of the status attainment research program has in fact been to ignore the existence of nonprestige dimensions of occupational differentiation and stratification. How could it

be otherwise? The transformation from occupational categories to prestige levels cannot be reversed. There is no way to transform occupational prestige scores into measures of relations to the means of production, or any other "structural" conception of occupational differentiation. Therefore it is difficult to interpret status attainment results in terms of alternative theoretical models of social structure or occupational differentiation. Similarly, there is no way for status attainment research to directly address issues of interest to those who do not accept the basic premise that occupational prestige is the only relevant aspect of occupational differentiation.

This is not a situation which is unique to the status attainment research. All theory-laden research areas will exhibit some such isolation from alternative theories. Indeed, Kuhn (1970) has argued that it is precisely such applied theoretical myopia which differentiates successful areas of scientific research from those which are less successful. In the present context I am more concerned with identifying the theoretical content of status attainment research than I am with evaluating that theoretical content. However, to the extent that such an identification requires a consideration of competing theories, some critical emphases may be unavoidable.

THE PROCESS OF STATUS ATTAINMENT

A major theme in the comments of status attainment critics and supporters alike is the concern with status attainment as a process model. What is the theory of social process implied by status attainment research? I have focused above on the measurement of occupations which define the origin and destination in the simple status attainment model. The concern with process requires attention to the set of intervening variables used to explain the transition from occupational origins to destinations. While there is a broad range of variables which have been included in status attainment models, the vast majority have been measured as and interpreted as individual characteristics. Be it parental expectations, IQ, educa-

tion, or personal aspirations, the variables in a status attainment model have been interpreted in terms of individual resources or liabilities which contribute to the individual attainment process.

Following Davis and Moore (1945), status attainment researchers are inclined to represent the process by which individuals are placed within the social and economic structure as that of recruitment of individuals to positions by way of the differential distribution of rewards across the positions. Duncan (1968:681-2), for example, identifies two processes which underlie stratification:

The first is the process of social metabolism, the recruitment of personnel to roles in a division of labor and the turnover of personnel in such roles. . . . The second process linked to differentiation is that of the allocation of rewards.

Such a representation serves to reinforce the individualist and voluntarist tendencies noted above by invoking a neo-classical, free market conception of occupational placement.

How can we explain the limited concern in the status attainment literature for extraindividual or structural constraints such as class barriers or between-group differences in opportunity structures? To do so, we must look to the theoretical heritage of functionalist theory in neo-classical conceptions of the socioeconomic order (Stolzenberg, 1975). Only if we assume an open, fully competitive market process in which individual characteristics are identified and rewarded according to their societal value can we justify ignoring market (structural) characteristics in the analysis of individual attainment. The competitive market situation assumed by neoclassical economic theories *guarantees* that the differential placement of individuals in the socioeconomic order is a reflection of the individual characteristics brought into the marketplace by the worker. Conversely, the interpretation of status attainment findings in terms of occupational attainment processes *requires* the assumption of market homogeneity for the population under study. (See Kerckhoff, 1976, for a discussion of problems and prospects of

nonindividualistic interpretations of status attainment research.)

METHOD OR THEORY

Status attainment findings, like any research findings, are conditional on the conceptual/theoretical assumptions imposed in undertaking analysis. While these assumptions have implications for the choice of methodological procedures to be employed in the analysis, they are not themselves methodological assumptions nor can they be justified in terms of methodological criteria. It is this point which Coser (1975) and Burawoy (1977) have failed to recognize, and it is this failure which serves to blunt the force of their criticisms of status attainment. Let us then consider how the characterization of status attainment as a theory-laden research tradition relates to such criticism.

Both Coser (1975) and Burawoy (1977) emphasize the homogeneous, individualistic orientation of the status attainment research in contrast to a preferred "structuralist" orientation:

The focus is predominantly on the impact on individual careers of differences in parental resources, access to educational institutions and the like, or they center attention upon individual characteristics of people variously placed in the class structure. There is no concern here with the ways in which differential class power and social advantage operate in predictable and routine ways, through specifiable social interactions between classes or interest groups, to give shape to determine social structures and to create differential life chances. (Coser, 1975:694)

These critics also concur in attributing the individualistic orientation to methodological considerations (Burawoy, 1977:1031; Coser, 1975:694) and contrasting this with their own theoretically-based concern with social structure. But with this juxtaposition several important questions are left unanswered. How is it that methods determine substance? In particular, why should methodological considerations lead to individualistic research? Will this always be the case, or is this a peculiarity of "linear statistics," as Burawoy (1977:1037) suggests?

The present exploration of the theory-

laden character of status attainment research suggests a different explanation for its individualistic orientation. Status attainment rests on a functionalist conception of social structure in which social positions are conceived of as levels of performance, which are differentially evaluated and rewarded within a competitive market situation. Given this conceptual framework, it is quite appropriate to represent occupational positions in terms of the evaluations of the general population:

Suffice to say here that people perceive rather accurately that professional and administrative occupations, by their very definition or organization, call for the exercise of greater authority and control and apparently require for their exercise, native and trained capacities and personality traits which craft or operative occupations, by their organization, do not (in degree or kind). (Featherman et al., 1975:333)

Similarly within the neoclassical theoretical heritage of functionalist theory, the assumption of fully open and competitive allocation of individuals to jobs (i.e., of market homogeneity) provides a source of justification for restricting attention to the individual characteristics of job-holders.

But once these basic theoretical assumptions have been used to structure the analysis—to decide what variables will be excluded from the analysis and how included variables will be measured—the choice of methodological procedures will have little impact on the basic focus and content of research. The homogeneity noted by critics derives from theoretical assumptions embedded in the research, not from the decision to employ regression analysis or some other procedure. Given these theoretical assumptions, stratification research becomes status attainment research, an inquiry into factors affecting the success/failure of individuals with different resources and abilities competing within an open opportunity structure. The basic orientation of status attainment research has nothing to do with quantitative methods, with path analysis, or with linear models. That orientation derives from the theory-laden character of status attainment research, not from the methodological procedures employed in that research.

One can interpret analyses of status attainment as analyses of social structure if and only if one assumes that occupational differentiation can best be represented as a single dimension of superiority/inferiority and that the allocation process by which individuals are placed in occupations is open and competitive across all individuals and occupations. But such a representation cannot be justified empirically. Status attainment proponents (Blau and Duncan, 1967; Klatzky and Hodge, 1971) and critics (Horan, 1974) alike have reported that intergenerational occupational mobility behavior cannot be reduced to a single (prestige/status) dimension of occupational differentiation. Likewise, analyses of the market structure of industrial societies are beginning to document the segmentation by which individuals from different social positions face different opportunity and reward structures (Bonacich, 1975; Beck et al., 1978a).

Thus the differences between the status attainment approach and a concern with class or structural factors as advocated by Coser and Burawoy is due not to methodology or to fact, but to theory. Status attainment research rests on a conception of occupational differentiation which denies the existence of the very aspects of occupational differentiation which are of essential interest from a structural orientation.

SOME IMPLICATIONS

The characterization of status attainment by its critics is important because it has implications for the research activities of those critics. If we attribute the theoretically unsatisfactory (from the perspective of critics) character of status attainment research to the use of quantitative methods, as Coser (1975) and Burawoy (1977) have suggested, then we are placed in the position of rejecting quantitative methods on the grounds that such methods lead to theoretically unsatisfactory results. This would tend to encourage sociologists with structural (as opposed to individualistic) orientations to avoid quantitative empirical research.

I have argued that this causal linkage

proposed by critics between methods and findings is nonexistent and that the individualistic character of status attainment research can be attributed to theoretical premises which are embedded in the basic conceptualization and measurement operations upon which the status attainment research program builds. This position has several implications for those favoring a structural orientation to stratification. First, structural researchers should realize the importance of linking *their* basic theoretical orientations to concrete conceptualization and measurement operations which can be employed in the analysis of survey data. Such activities are a prerequisite to the growth of a useful, nonindividualistic research program in social structure and social stratification. Second, structural theorists and researchers should learn to approach the results of status attainment research with considerable caution, and to recognize the extent to which such results are conditional upon theoretical assumptions antithetical to their own theoretical orientations.

While status attainment research may be expected to continue to hold a dominant position in American stratification research for some time, there are several promising alternatives for researchers who do not accept the theoretical premises of the status attainment model. One alternative is the development and use of a neo-Marxist classification of occupational data—an old ideal given new impetus by the work of Wright and Perrone (1977). These authors argue that the conventional occupational information available from most survey data is an inadequate basis for identifying Marxist class positions—i.e., “positions within the social relations of production” (Wright and Perrone, 1977:35). Using information on ownership of means of production, purchase/sale of labor power, and control of labor power, they distinguish four class categories (employers, managers, workers, petty bourgeoisie) and examine the effects of the first three on income attainment and inequality. Their analysis identifies not only substantial net class effects on income but also some interesting class effects on the income returns to education.

Another promising alternative to the status attainment research tradition derives from the dual economy literature (Averitt, 1968; Bluestone et al., 1973). This literature introduces the concept of economic sectors—structural entities which derive from the nature of modern industrial capitalism—as an important factor in the income determination process. Distinguishing between a core industrial sector dominated by monopolistic capitalist firms and a periphery sector characterized by competitive capitalism, the dual economy literature suggests that the sectoral placement of a worker may condition the income returns to individual characteristics such as education. Bibb and Form (1977) demonstrate that for blue-collar workers, industrial sector and other structural characteristics such as firm size, occupational skill level, and SMSA residence provide substantial increments in income variance-accounted-for over models including only individual characteristics. Beck et al. (1978a) find that sectoral differences in earnings cannot be explained away by differential labor force composition and that the financial returns to individual characteristics such as schooling, sex, race and age are substantially different in the core and periphery sectors.

These new directions in American stratification research are important for several reasons. First, they provide negative empirical evidence concerning certain critical assumptions of status attainment research: the adequacy of occupational prestige as a measure of social position and the adequacy of social process models which include only individual characteristics. Second, each provides a nonindividualistic theoretical foundation for the quantitative analysis of survey data, and uses the same basic set of methodological procedures (regression and covariance analysis) as status attainment research. Third, each of these new approaches to stratification research has important implications both for the analysis of mobility and for the analysis of racial and sexual discrimination in the economic sphere. The pursuit of these implications will be an important focus for research deriving from both the neo-

Marxist and the dual economy theoretical perspectives during the next decade (see, e.g., Beck et al., 1978b).

REFERENCES

- Averitt, Robert T.
1968 *The Dual Economy: The Dynamics of American Industry Structure*. New York: Horton.
- Beck, E. M., Patrick M. Horan, and Charles M. Tolbert II
1978a "Stratification in a dual economy: a sectoral model of earning determination." *American Sociological Review*. In press.
1978b "Labor market segmentation and discrimination against minorities." Paper to be presented at the annual meeting of the American Sociological Association, San Francisco.
- Bibb, R. and W. Form
1977 "The effects of industrial, occupational and sex stratification on wages in blue-collar markets." *Social Forces* 55:974-96.
- Blau, Peter M. and Otis Dudley Duncan
1967 *The American Occupational Structure*. New York: Wiley.
- Bluestone, Barry, William M. Murphy and Mary Stevenson
1973 *Low Wages and the Working Poor*. Institute of Labor and Industrial Relations, University of Michigan, Ann Arbor.
- Bonacich, Edna
1975 "Advanced capitalism and black/white relations in the United States: a split labor market interpretation." *American Sociological Review* 41:34-51.
- Burawoy, Michael
1977 "Social structure, homogenization and the process of status attainment in the United States and Great Britain." *American Journal of Sociology* 82:1031-42.
- Coser, Lewis A.
1975 "Presidential address: two methods in search of a substance." *American Sociological Review* 40:691-700.
- Davis, K., and W. E. Moore
1945 "Some principles of stratification." *American Sociological Review* 10:242-9.
- Duncan, Otis Dudley
1961 "A socioeconomic index for all occupations." Pp. 109-38 in A. J. Reiss (ed.), *Occupations and Social Status*. New York: Free Press.
1968 "Social stratification and social mobility: problems in the measurement of trend." Pp. 675-719 in Eleanor Bernert Sheldon and Wilbert E. Moore (eds.), *Indicators of Social Change*. New York: Russell Sage.
- Duncan, Otis Dudley and Robert W. Hodge
1963 "Education and occupational mobility." *American Journal of Sociology* 67:629-44.
- Featherman, D. L., F. L. Jones, and R. M. Hauser
1975 "Assumptions of social mobility research in the U.S.: the case of occupational status." *Social Science Research* 4:329-60.

- Goldthorpe, J. H. and K. Hope
1964 *The Social Grading of Occupations: A New Approach and Scale*. Clarendon: Oxford University Press.
- Gordon, David M.
1972 *Theories of Poverty and Underemployment*. Lexington: Heath.
- Hanson, N. R.
1958 *Patterns of Discovery*. Cambridge, Ma.: Cambridge University Press.
- Horan, Patrick M.
1974 "The structure of occupational mobility: conceptualization and analysis." *Social Forces* 53:33-45.
- Kerckhoff, Alan C.
1976 "The status attainment process: socialization or allocation?" *Social Forces* 55:368-81.
- Klatzky, Sheila R. and Robert W. Hodge
1971 "A canonical correlation analysis of occupational mobility." *Journal of the American Statistical Association* 66:16-22.
- Kuhn, Thomas S.
1970 *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Parsons, Talcott
1940 "An analytic approach to the theory of social stratification." *American Journal of Sociology* 45:841-62.
- Reiss, Albert J.
1961 *Occupations and Social Status*. New York: Free Press.
- Siegel, P. M.
1971 *Prestige in the American Occupational Structure*. Ph.D. dissertation, Department of Sociology, University of Chicago.
- Stolzenberg, Ross M.
1975 "Occupations, labor markets and the process of wage attainment." *American Sociological Review* 40:645-65.
- Wright, Erik Olin and Luca Perrone
1977 "Marxist class categories and income inequality." *American Sociological Review* 42:32-55.

NONRECURSIVE MODELS OF LABOR FORCE PARTICIPATION, FERTILITY BEHAVIOR AND SEX ROLE ATTITUDES*

LYNN SMITH-LOVIN AND ANN R. TICKAMYER

University of North Carolina

American Sociological Review 1978, Vol. 43 (August):541-557

Studies of female labor force participation and fertility have established a negative relationship between the two; however, there has been much debate over the direction and causes of this relationship. In this paper, two causal models of the actual fertility and work behavior of a national sample of married women, aged 30 in 1970, are examined using the two-stage least-squares technique to disentangle reciprocal effects. Included are a two-variable feedback loop incorporating only fertility and labor force participation and a three-variable model which adds sex role attitudes to the endogenous variables. Most of the work-fertility relationship can be accounted for by controlling background variables such as education and marital duration. A negative effect from fertility to labor force participation remains, however. Adding sex role attitudes to the model as a potential source and consequence of fertility and work behavior slightly reduces the size of this effect.

General reviews of the microlevel research on fertility and female labor force participation usually agree that there is an inverse relationship between childbearing and work outside the home (Germain, 1975; Piepmeyer and Adkins, 1973; Lowenthal and David, 1972). There is less

agreement, however, on the direction of causal influence. Depending upon the interests and initial assumptions of the investigators, either work or fertility has been used as the dependent variable and the findings are presented accordingly (Sastry, 1975; Bumpass and Westoff, 1970; Ridley, 1959; 1969; Pratt and Whelpton, 1958).

Understanding this relationship is important for two distinct but related reasons: for more extensive knowledge of basic demographic processes and for information necessary for policy formula-

* We are grateful to Bruce K. Eckland, David R. Heise, Michael Hout, Ronald R. Rindfuss and two anonymous reviewers for comments on an earlier version of this paper. Portions of this paper were presented at the annual meeting of the Population Association of America, St. Louis, 1977. The authors' names are listed in alphabetical order to indicate equal contribution to this research.

tion in a world concerned with population growth. The latter goal has received particular attention because of its potential for discovering variables which can be manipulated to bring about desired policy outcomes. For example, a number of authors have suggested that birth rates can be lowered by altering the labor force structure to increase women's labor force participation (Blake, 1970; Hoffman and Nye, 1974). This view rests on the assumption that the direction of causal influence is from women's work status to fertility plans and behavior. However, this has yet to be established empirically. Therefore, it is necessary to understand more fully the basic social and demographic processes involved.

A number of different explanations have been suggested to account for this inverse relationship between work and fertility. Most prominent are role compatibility formulations and microeconomic theory. In the former, it is hypothesized that in a modern industrialized economy, mother and worker roles are incompatible and, therefore, women who jointly occupy these roles will have fewer children (Stycos and Weller, 1967). Reasons for this incompatibility are a bureaucratized occupational structure, which does not permit the flexibility required for childbearing, and nuclear families, which, in the absence of other forms of child care, leave women with no alternatives but to assume the entire burden of childrearing. Microeconomic models suggest that under similar conditions, the utility of children decreases along with a rise in the costs of childbearing and the opportunity costs for mothers (Becker, 1976; Mincer, 1963; Turchi, 1975; Easterlin, 1973). When child labor is no longer a valued household commodity, and the costs to the family of each child skyrockets, then couples will decide to have fewer children. Particularly important is the assumption that women increasingly become aware of the opportunity costs childrearing has for them (Stolzenberg and Waite, 1977). With expanded opportunities for women to gain income, prestige, and psychic gratification outside the family, childbearing becomes an alternative which exacts a heavy price.

Both of these approaches usually are

interpreted to imply that where control of fertility is possible, childbearing is affected greatly by work decisions. However, a number of other studies have posited the reverse—that fertility determines labor participation (Sweet, 1973; Bowen and Finnegan, 1969; Cain, 1966). Although the reasoning behind this hypothesis is not as well-developed, the implication is that childbearing is the activity of choice for young women, and everything else is subsumed to this task. Pressures from the normative timing of life course events mandate early adulthood as a period of family building. The vast majority of young married women intend to have children and carry out these intentions during the early years of marriage. Given the priority of traditional roles, extrafamilial activities will be arranged around them and possibly will be dropped entirely to accommodate their demands. Normative expectations about childrearing for women traditionally have required that precedence be given to the parental role, even to the extent of suggesting that a mother's labor force participation is harmful to young children's development. Furthermore, the constraints imposed by the occupational structure, the lack of good opportunities for women in the labor force, and the inflexibility of young children's schedules make working a less preferred and extremely difficult activity for women. In summary, according to this view, married women are more likely to work if they have fewer children.

There are yet other interpretations of the labor force participation-fertility relationship. These are that it is a spurious relationship (i.e., both fertility and labor force participation are the result of other factors) and that there is reciprocal causation (Blake, 1970; Terry, 1974; Waite and Stolzenberg, 1976).

It is the last hypothesis that recently has seemed most theoretically compelling. It is assumed that women who work will have fewer children for all the reasons mentioned earlier and, conversely, having fewer children will enable women to work. However, it is only in the past few years that sociologists have become aware of techniques permitting them to test empirically such complex relationships.

Waite and Stolzenberg (1976), for example, use two-stage least-squares analysis to demonstrate that there is simultaneous influence of labor force participation plans on fertility plans, and vice versa. They find that while both effects are negative, the work to fertility path is much stronger than the effect from fertility plans to labor force plans. Although the accuracy of such results depends upon the validity of the assumptions used to specify the model, the use of such econometric techniques to analyze non-recursive models of fertility and women's work seems to hold promise for the future.

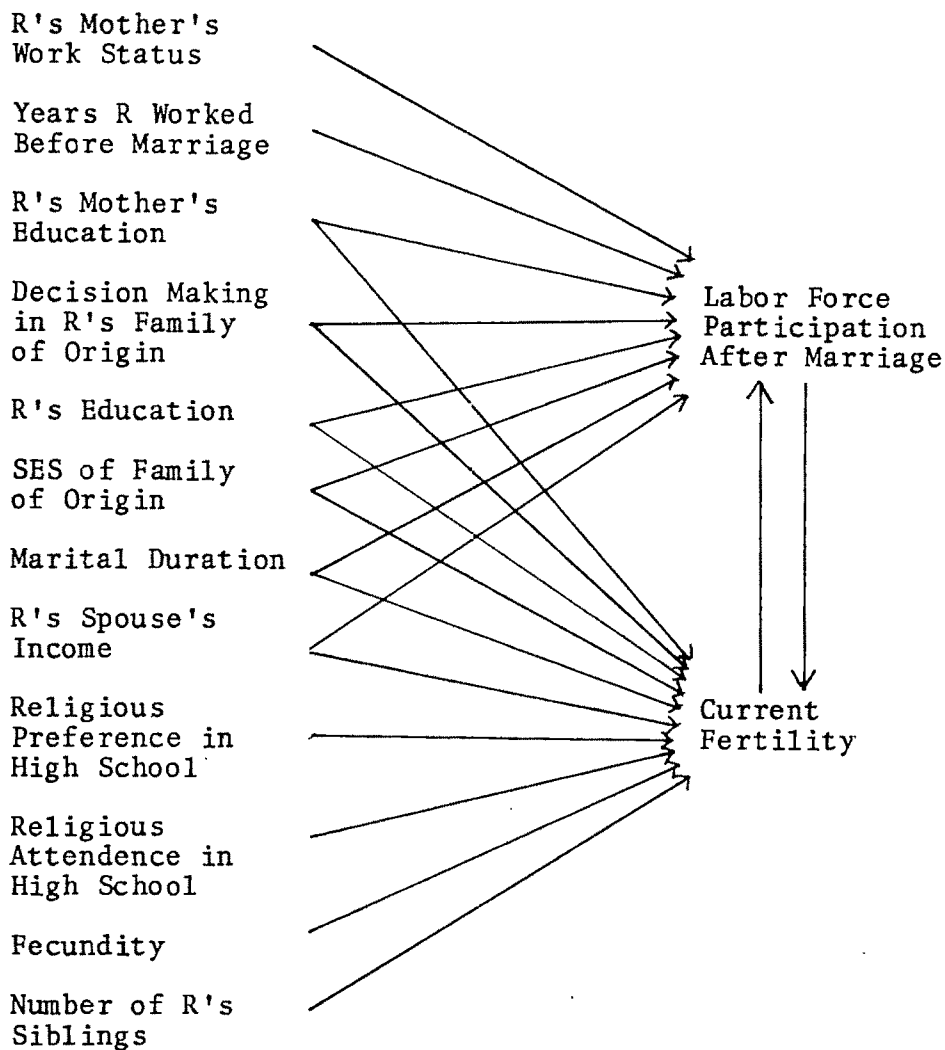
This paper studies actual fertility and work *behavior* rather than fertility and work *plans*. Although analyses showing the relationships among various plans and/or expectations are interesting, most of the theoretical work done in the fertility-labor force participation area has concentrated on the incompatibility of the behaviors entailed by mother and worker roles (for example, Boulding, 1976; Stycos and Weller, 1967). To the extent that young women perceive such incompatibility, one would expect fertility and work plans to be related negatively. But the causal relations between such planning variables need not be entirely consistent with the links between behavior variables. Especially during the early years of marriage, childbearing and work activities may influence each other (and be influenced by other factors) in ways that young respondents might not anticipate. Therefore, this paper will focus on the actual experiences of married women interviewed in their thirties. Two separate models of the work-fertility relationship are examined. One is similar to the Waite and Stolzenberg model in that it examines a potentially reciprocal relationship between fertility and female labor force participation. However, instead of looking at work and fertility plans, this model has actual fertility¹ and labor force behaviors as dependent variables (see Figure 1).

¹ Since the respondents were still in their childbearing years at the time of the survey, the effects of other variables on fertility could be due to either changes in the number of children a woman has or changes in the timing of those children. Therefore, some analyses were also conducted with ex-

The second model proposed here differs in another respect from that developed by Waite and Stolzenberg. Sex role attitudes are included as both a potential cause of work and fertility behaviors and a potential consequence of having engaged in these activities to varying degrees. To estimate the parameters of their model, Waite and Stolzenberg make the assumption that work attitudes affect labor force participation plans but have no direct effect on fertility expectations. In addition, they must assume that the work attitudes variable is not correlated with the unspecified sources (i.e., the error term) of the fertility expectations variable (Heise, 1975). The theoretical argument used to support these assumptions notes that the attitudinal measure deals only with the costs and benefits of work for married women. But if young women are aware of the potential incompatibility of the worker and mother roles, their attitudes toward childbearing and their more general sex role conceptions almost certainly will be correlated with the work attitudes variable. Such fertility values and sex role attitudes would also affect fertility expectations. If so, these unspecified sources would then violate the assumptions of the Waite and Stolzenberg two-stage least-squares analyses.

In our second model of fertility and work behaviors (presented in Figure 2) sex role attitudes are conceptualized as stable orientations toward the worker and mother roles and the trade-offs which must be made between them. Such orientations are an important potential source of fertility and work behavior during young adulthood; they also, of course, may be influenced by fertility and work experiences. By including sex role attitudes explicitly in our model, the theoretically important relationship between these predispositions and behaviors can be

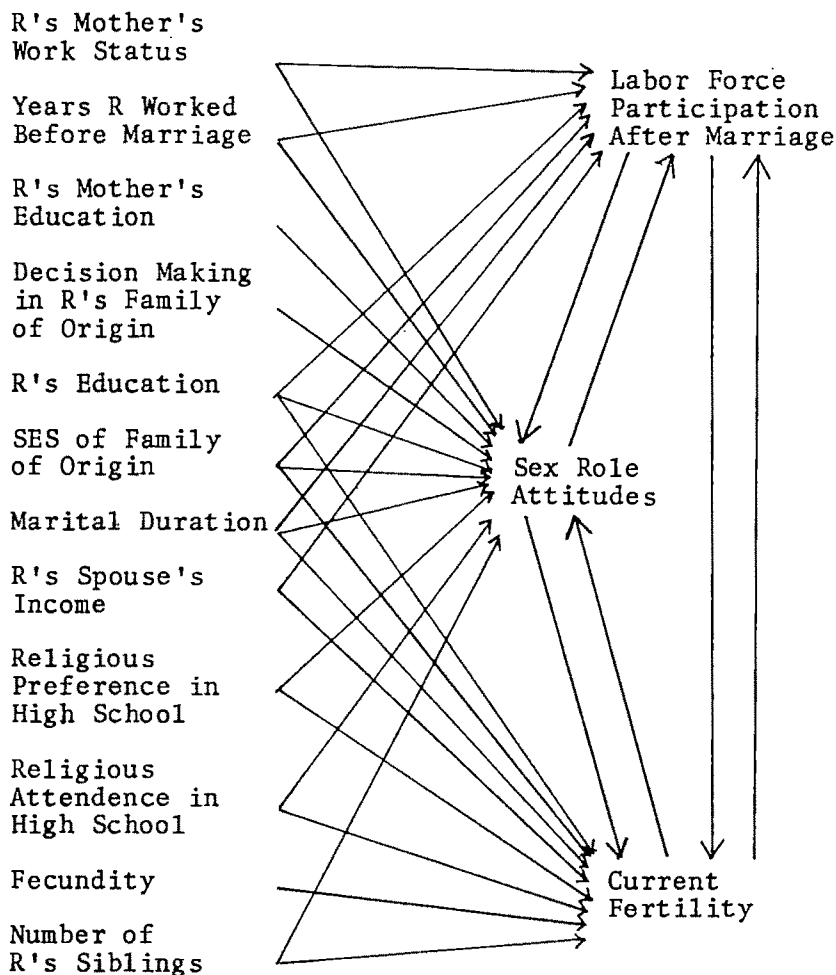
pected fertility (current fertility + number of additional children expected) as the dependent variable. If the effects of a variable on current fertility are due primarily to its effects on the timing of births (early vs. late), that effect should disappear when expected family size is analyzed. If the effects are due to differences in completed fertility, the relationships should be similar in the analyses using current and expected fertility measures.



$$\begin{aligned} \text{LFP} = & b_0 + b_1 \text{ FERTILITY} + b_2 \text{ WORKMO} + b_3 \text{ YRSWORKBM} + \\ & b_4 \text{ EDUCMO} + b_5 \text{ DECMARK} + b_6 \text{ EDUCR} + b_7 \text{ FAMSES} + \\ & b_8 \text{ YRSMAR} + b_9 \text{ SPINC} + U \end{aligned}$$

$$\begin{aligned} \text{FERTILITY} = & c_0 + c_1 \text{ LFP} + c_2 \text{ EDUCMO} + c_3 \text{ DECMARK} + c_4 \text{ EDUCR} \\ & + c_5 \text{ FAMSES} + c_6 \text{ YRSMAR} + c_7 \text{ SPINC} + c_8 \text{ RELGOR} + \\ & c_9 \text{ RELGFRHS} + c_{10} \text{ FECUND} + c_{11} \text{ NSIBS} + V \end{aligned}$$

Figure 1. A Nonrecursive Model of Labor Force Participation and Fertility Behavior for a Cohort of Married Women



$$\text{LFP} = b_0 + b_1 \text{ FERTILITY} + b_2 \text{ SEXROLE} + b_3 \text{ WORKMO} + b_4 \text{ YRSWORKBM} \\ + b_5 \text{ EDUCR} + b_6 \text{ FAMSES} + b_7 \text{ YRSMAR} + b_8 \text{ SPINC} + U$$

$$\text{FERTILITY} = c_0 + c_1 \text{ LFP} + c_2 \text{ SEXROLE} + c_3 \text{ EDUCR} + c_4 \text{ FAMSES} + \\ c_5 \text{ YRSMAR} + c_6 \text{ SPINC} + c_7 \text{ RELGOR} + c_8 \text{ RELGFRHS} + \\ c_9 \text{ FECUND} + c_{10} \text{ NSIBS} + V$$

$$\text{SEXROLE} = d_0 + d_1 \text{ LFP} + d_2 \text{ FERTILITY} + d_3 \text{ WORKMO} + d_4 \text{ YRSWORKBM} \\ + d_5 \text{ EDUCMO} + d_6 \text{ DECMAKE} + d_7 \text{ EDUCR} + d_8 \text{ FAMSES} + \\ d_9 \text{ YRSMAR} + d_{10} \text{ RELGOR} + d_{11} \text{ RELGFRHS} + d_{12} \text{ NSIBS} + W$$

Figure 2. A Nonrecursive Model of Labor Force Participation, Fertility Behavior, and Sex Role Attitudes

examined and this source of possible bias in the estimates of the causal effects of work and fertility may be removed.

One final consideration guides the development of these models. Although studies typically have related current work status to current fertility, the literature on the measurement of behavior variables suggests that the relationships will be strengthened if multi-item scales are used (Fishbein and Ajzen, 1974; Tittle and Hill, 1967). Measures are needed which capture overall behavioral patterns. Fertility is already the cumulative product of a series of behaviors (Namboodiri, 1972). Labor force participation, however, is more of a problem. Since fertility represents a reproductive process which has developed throughout a woman's life, it is appropriate to construct a work measure which represents labor force participation throughout the same period. Therefore, a variable which measures women's work status at many points in time during their married life is needed. One possibility for such a measure is the proportion of married life spent in the labor force (i.e., years worked/years married) (Groat et al., 1976). However, such ratio variables have undesirable measurement characteristics when used for microlevel analyses (Bohrnstedt and Marwell, 1977). Therefore, in this analysis the women's work behavior during their married life is represented by the number of years employed since marriage. The length of the interval for potential employment (as determined by marital duration and the interval since completion of schooling) is controlled by other variables in the model.²

² Educational attainment and age at marriage (which is related very closely to marital duration in an age cohort) combine to determine the interval that a woman has available for employment both before and after marriage. The more years a woman spends in school, the fewer she will have available for employment; the earlier she marries, the fewer years she will have available for work before marriage and the more years she will have available for work after marriage. Since these effects are linear and additive, having educational attainment and marital duration in the model as exogenous control variables corrects for the length of the intervals available for work before and after marriage (in data from an age cohort) and makes the use of a proportion as the dependent variable unnecessary. (To check this rea-

The use of nonrecursive models entails special problems for data analysis. In this paper a discussion of identification problems and their possible solutions for both Models 1 and 2 will be presented, followed by descriptions of the data and the measurement procedures used to estimate the models' parameters.

SPECIFICATION OF THE MODEL

Both Models 1 and 2 have endogenous variables in feedback loops. To estimate the parameters of these nonrecursive causal relationships, instrumental variables which meet the following theoretical requirements are needed: a variable X is an instrument for Y in the $Y \rightarrow Z$ relationship if it either causes Y or is correlated with Y but does not directly cause Z . Furthermore, X cannot be caused by either Y or Z , and it cannot be correlated with the unspecified sources of Z (Duncan, 1975; Goldberger, 1964; Heise, 1975). For example, in Model 1 the couple's fecundity may serve as an instrument for fertility in the fertility \rightarrow labor force participation relationship. Fecundity obviously affects fertility, but it should not have any causal impact on work behavior except through the fertility variable. Furthermore, since fecundity is determined by nonvoluntary physiological factors (contraceptive sterilizations are coded as "fecund"), it should not be affected by work participation. Similarly, it is not expected to be correlated with any unspecified sources of the loop variables. In Model 2 fecundity serves as an instrument for both the fertility \rightarrow work relationship and the fertility \rightarrow sex role attitude relationship.

soning, the analysis also was carried out with the proportion as the dependent variable. None of the coefficients shift markedly, the significance of the coefficients remains unchanged, and substantive conclusions would be the same.) One additional point should be emphasized. Using marital duration and educational attainment as control variables does not allow investigation of the very complex process by which these factors are determined—the relationship between education, age at marriage, and work before marriage is represented by unanalyzed exogenous intercorrelations. However, including them in the model does allow estimation of the independent effect of an increase in education, for example, on work after marriage.

At least one instrumental variable must be found for each nonrecursive relationship. These instruments are listed below for each model.

Model 1

Variables affecting only labor force participation (instruments for LFP → FERTILITY): mother's work status while R was growing up and the number of years R worked before she married. R's mother's work history should influence R's participation in the labor force by providing a positive role model of a working woman. Since R's fertility and work occur years after her mother's behavior, neither of these variables can have a return effect on mother's work status. Also, the mother's work behavior is unlikely to be correlated with unspecified sources of R's fertility. Although the mother's work may have influenced her own fertility behavior, the possible effect of this family size role model should be controlled by including the number of R's siblings as a variable in the model. R's work experience prior to marriage is expected to influence her later work behavior by increasing both her employability and her earning potential. Since this work experience occurs before the period measured by the dependent variables, they cannot have any causal impact on the earlier behavior.³ To be used as an instrument, R's work prior to marriage also must be uncorrelated with any unspecified sources of fertility and work during marriage. This may be an unjustified assumption, since R's work experience before marriage might well be associated with more general sex role orientations which are probably a source of both fertility and later labor force participation. Model 2 is an attempt to resolve this problem.

Variables affecting only fertility (instruments for FERTILITY → LFP): the couple's religious origins (i.e., whether one or both were Catholic in high school),

frequency of R's religious participation in high school, the couple's fecundity, the number of R's siblings. Catholic upbringing is known to lead to higher fertility (Bumpass and Westoff, 1970; Ryder and Westoff, 1971), but the temporal priority of this variable to fertility and work behavior prevents any reciprocal causation. To be used as an instrument, we also must assume that religious origins are uncorrelated with unspecified sources of the fertility and work variables. (The same objections noted above could challenge the validity of this assumption. Again, Model 2 will attempt to deal directly with this problem.) Frequent participation in religious services during the teenage years is included in the model as an indicator of a more traditional religious upbringing. Since almost all faiths stress the importance of family life, greater exposure to religious doctrine is expected to influence fertility. As with the religious preference variable, temporal priority precludes any effect of the loop variable on religious upbringing. Similarly, the assumption that frequency of religious participation is uncorrelated with unspecified sources of the loop variables is suspect. Fecundity, the biological capacity to have children, obviously influences fertility, but cannot be affected by fertility and work. It is unlikely to directly affect labor force participation or to be correlated with unspecified sources. The number of R's siblings is included as a variable in the model primarily because Waite and Stolzenberg (1976) argue that it is a suitable instrument for the fertility → work relationship. They suggest that a large number of brothers and sisters creates a "taste for larger families."

Model 1 also includes six background variables which affect both endogenous variables: spouse's income, R's mother's decision-making power in her family of origin, the socioeconomic status of R's family of origin, R's educational attainment, R's marital duration, and the educational attainment of R's mother.

Our second model includes sex role attitudes as a potential source of fertility and work behaviors and a potential result of fertility and work experiences. The introduction of a third endogenous variable

³ This assumption ignores the potential impact of premarital conceptions on age at marriage and therefore on work before marriage. Since such conceptions appear to be quite uncommon in this sample of women (Eckland and Fried, 1975), this factor should not seriously bias parameter estimates.

greatly complicates the model, but it is necessary to clarify some of the problematic assumptions of Model 1. For example, to be used as an instrument for the $LFP \rightarrow FERTILITY$ relationship, R's mother's work status and R's work experience before marriage must have no direct effect on fertility and must be uncorrelated with unspecified sources of fertility. These assumptions are likely to be incorrect if sex role attitudes are not in the model. Such factors might easily influence sex role conceptions, which are likely to be an unspecified source of fertility. Similarly, several of the instruments for $FERTILITY$ in Model 1 (religious origins, religious participation, and number of siblings) might be correlated with sex role attitudes. If such attitudes are an unspecified source of work behavior, then this correlation violates the assumptions necessary for the use of these variables as instruments. All of these potential specification problems can be solved by including sex role attitudes explicitly in the model. However, adding this endogenous variable also creates new nonrecursive relationships which must be identified. Below, instrumental variables for each of the six nonrecursive effects in Model 2 are presented.

Model 2

Variables affecting only fertility (instruments for $FERTILITY \rightarrow LFP$ and for $FERTILITY \rightarrow SEXROLE$): the couple's fecundity. The same arguments presented in the discussion of Model 1 apply here.

Variables affecting only sex role attitudes (instruments for $SEXROLE \rightarrow FERTILITY$ and for $SEXROLE \rightarrow LFP$): R's mother's decision-making power in the family of origin and R's mother's educational attainment. Both of these variables are expected to influence R's sex role attitudes by providing her with a positive role model for nontraditional sex roles while she was growing up. Since these characteristics of R's family background cannot directly increase her earning potential or employability after her marriage, these variables can only influence labor force participation through attitudes and values which were formed in the earlier period.

Similarly, we expect any effect on fertility to be mediated by sex role orientations. Prior occurrence prevents any feedback effects from the loop variables.

Variables affecting only fertility and sex role attitudes (instruments for $FERTILITY \rightarrow LFP$ and for $SEXROLE \rightarrow LFP$): the couple's religious origins, R's religious participation in high school, and the number of R's siblings. As noted in the discussion of Model 1, religious upbringing is expected to affect both sex role attitudes and family size orientations. By including sex role attitudes explicitly in Model 2, the possibility that religious background variables will be correlated with unspecified sources of the labor force participation variable is reduced considerably. Also, since the size of R's family of origin might influence both her role models and her family size "tastes," number of siblings is specified as a potential source of sex role attitudes as well as fertility.

Variables affecting only fertility and labor force participation (instruments for $FERTILITY \rightarrow SEXROLE$ and for $LFP \rightarrow SEXROLE$): R's spouse's income. Income is a well-documented influence on both fertility and female labor force participation (Rosenberg, 1973; Simon, 1974), but since work is a normative activity for adult men we do not expect these variables to influence the spouse's labor force behavior in return. Furthermore, there is no reason why husband's income should influence the sex role attitudes of the woman directly.

Variables affecting only sex role attitudes and labor force participation (instruments for $SEXROLE \rightarrow FERTILITY$ and for $LFP \rightarrow FERTILITY$): R's mother's work status when R was young, and the number of years R worked before marriage. The same arguments presented in the discussion of Model 1 apply here as well. A positive role model and work experience would influence sex role conceptions. In addition, work experience prior to marriage would increase R's employability and earning potential in later years.

Finally, several background factors are expected to influence all three endogenous variables. These are the socioeco-

conomic status of R's family of origin, R's educational attainment, and marital duration.

DATA

The data set used was provided by the 1970 EEO Survey ("Explorations in Equality of Opportunity") conducted by the Institute for Research in Social Science, University of North Carolina. These data represent a 15-year follow-up survey of a national sample of high school youth who originally were surveyed by the Educational Testing Service in 1955, when they were sophomores in high school.

Only data from the 1970 follow-up survey will be used in the present analysis. The basic sampling units were schools, which were stratified by region, size, level of parental education, group test scores, school drop-out rates and college-going rates. The former students were sent a 36-page questionnaire covering family, educational and career histories. Four mailings and personal contact resulted in a response rate (based upon the number of persons for whom addresses were available) of 60%. Overall, data were obtained for 2,077 of the sample sophomores, 1,126 females and 951 males.⁴

It should be emphasized that the sample does represent a single age cohort; all respondents were approximately 30 at the time of the 1970 follow-up. This is both

an asset and a liability. Age and historical experience are controlled, but since age is related to a number of variables in the analysis, including sex role attitudes and employment status, the relationships found may hold only for the particular cohort studied.

For the present analysis, the sample was restricted to include only women who were still married to their first husband at the time of the survey. This was desirable both to eliminate the complicating effects of disrupted marital histories and to control for risk. Women who had not yet completed their education also were excluded in order to eliminate possible feedback effects from the dependent variables to educational attainment. Respondents with missing data on the dependent variables also were deleted resulting in a final N of 703.

Respondents with missing data on the exogenous variables (which averaged 1-3% with the exception of spouse's income which was somewhat higher) were assigned the median value if the variable was interval or ordinal and were assigned the mode if the variable was a nominal dichotomy. This is a conservative procedure which tends to attenuate slightly the relationships among the model variables.

MEASUREMENT

Exogenous Variables

The exogenous variables are the same for both models and therefore the procedures for measuring them will be described simultaneously.

R's mother's work status (WORKMO). The respondent's mother's work status while R was in high school was measured with a dichotomous variable which was coded one if the mother was employed and zero if the mother did not work.⁵

⁴ Although the response rate on the EEO data set is lower than desirable, it is one of the few available large national samples which contains information on all of the variables needed to identify these models, and, therefore, to investigate the nonrecursive relationships among the endogenous variables. Several data sets exist which have all the variables except fecundity, but this information is essential for the identification of the fertility → work relationship. The drawbacks of the data set are mitigated by the fact that it has been used to investigate a wide variety of topics, and, therefore, the characteristics of the sample are quite well-known. Furthermore, Alexander and Eckland (1973;1975) studied the possible response bias in this data and concluded that although the respondents underrepresent students from large urban areas and low-aptitude students, estimates of relationships among variables in the data do not seem to be affected. Comparisons of female respondents and nonrespondents showed that they do not differ significantly in marital status, number of children or migration status.

⁵ The type of work in which R's mother was engaged may have a major impact on the type of role model she presented to her daughter. To check the possibility of such effects, the analysis also was conducted substituting two dummy variables, ROLE1 and ROLE2, for the WORKMO variable. ROLE1 was coded 1 if R's mother worked as either an unskilled worker, a semiskilled worker, a blue-collar worker, a service worker, a saleswoman, a retail

Years R worked before marriage (YRWORKBM). Information for this measure was gathered from detailed life history tables which included both family and work data. The number of years that the respondent reported working before the year she reported getting married was computed.

R's mother's education (EDUCMO). This is the respondent's report of the number of years of schooling that her mother completed.

Decision making in R's family of origin (DECMAKE). This variable was constructed from four highly correlated items which asked, "To what extent would each of the following statements have applied to your mother and to your father when you were younger? Made important family decisions, Handled finances, . . ." The answers ranged from "Strongly" to "Not at all" for each parent. Answers referring to the father were recoded so that higher scores indicate the mother's greater responsibility for family decision making for each item. The items were summed to form the final measure.

R's educational attainment (EDUCR). This is the respondent's report of how many years of schooling she completed.

Socioeconomic status of the family of origin (FAMSES). A measure of the socioeconomic status of R's family of origin was formed by combining measures of the father's occupational status, the father's educational attainment and R's subjective judgment of how wealthy her family was in relation to the other children's families at her high school. An absolute measure of wealth was not used because a large proportion of the respondents had missing data on these items.

clerk, or an office or clerical worker; it was coded 0 if she was not employed in one of these categories. ROLE2 was coded 1 if the respondent's mother was a technical worker, a manager or owner of a business, or a professional; it was coded 0 if she did not fall into one of these categories. In effect, the two dummy variables compare respondents whose mothers worked at low-status jobs and those whose mothers worked at high-status jobs to those whose mothers did not work at all (the deleted category). The ROLE1 and ROLE2 variables had no significant effects on sex role attitudes, fertility or work behavior.

Marital duration (YRSMAR). This measure was constructed from the same life history tables mentioned above. The years which the respondent reported she was married were summed to form the variable. Since this data represents an age cohort, marital duration is a function of age at marriage. Therefore, a number of different factors may be confounded in this variable. It is introduced primarily as a control.

R's spouse's income (SPINC). This variable was measured by subtracting R's earnings from the total household income for the preceding year. The objective was to obtain a measure of the family's general financial situation independent of the income contributed by the wife's employment. The items measuring income had 24 categories each.

Religious preference in high school (RELGOR). If both the respondent and her husband were raised as Catholics, this variable was coded two. If either one but not both were raised as a Catholic, it was coded one. If neither member of the couple had a Catholic upbringing, it was coded zero.

Religious attendance in high school (RELGFRHS). This represents a measure of how often the respondent reports attending religious services while in high school. Responses ranged from "Not at all" (coded zero) to "Greater than once a week" (coded five).

Fecundity (FECUND). This measure was developed from the respondent's self-report of whether she or her husband were unable to have a(nother) child. Those who said that they were capable of further childbearing were coded one. Respondents who reported that either they or their husband had undergone a contraceptive sterilization also were classified as fecund (and coded one), since the variable sought to measure involuntary subfecundity which could not be affected by other variables in the model (e.g., current fertility). All others (i.e., those who reported that they could not bear children and who had not had a voluntary sterilization) were coded zero.

Number of R's siblings (NSIBS). This was the total number of brothers and sisters the respondent reported.

Endogenous Variables

Both Models 1 and 2 have the following two endogenous variables.

Current fertility (FERTILITY). The number of children that the respondent reports having at the time of the survey was coded as current fertility.

Years of married life spent in the labor force (LFP). Information for this measure was gathered from the detailed life history tables mentioned above. This variable represents the number of years that the respondent reported working, beginning with the year she reported her marriage.⁶

Additionally, Model 2 includes a third endogenous variable.

Sex role attitudes (SEXROLE). A scale was formed by factor analyzing 24 Likert-type attitude items relating to women's roles. The items were coded from one to four with one indicating strong agreement with a modern sex role orientation and four indicating strong disagreement. The final scale was computed by summing the 15 items which loaded highly on a factor which represented attitudes toward mothering and careers as proper women's roles.⁷

⁶ Part-time work while still in school was not considered labor force participation for the purpose of coding this variable. Part-time work for pay after the completion of schooling was coded as labor force participation. This type of activity was relatively rare in our sample.

⁷ The 15 items included in the scale measuring sex role attitudes and their factor loadings are as follows:

<i>Item</i>	<i>Factor Loading</i>
*1. Although women hold many important jobs, their proper place is in the home.	.491
2. A woman can live a full and happy life without marrying.	.496
*3. Career women tend to be masculine and domineering.	.390
*4. Children make a marriage happy.	.570
*5. Women should have as many children as God gives.	.423
*6. It is more important for a woman to help her husband than to have a career herself.	.351
7. There are times when children are really not very rewarding.	.429
*8. Women should choose only career fields that can help them be better mothers.	.398

RESULTS

Two-stage least-squares analysis was used to obtain parameter estimates for the models.⁸ The results of the analyses are presented in Tables 1 and 2.

Model 1

The strongest determinants of labor force participation in this model are R's education, work experience prior to marriage, marital duration, spouse's income and current fertility. Education, work experience and marital duration all have positive effects on the proportion of married life spent in the labor force, indicating that factors affecting employability and earning potential increase work participation, as does length of marriage. Spouse's income has a negative effect, indicating economic pressure to work for women whose husbands have low incomes. The effect of fertility on labor force participation is negative as expected. No other variables in the model have effects which are substantively or statistically significant.

Only three variables have significant effects on fertility. These are fecundity, marital duration and religious preference in high school. Couples who have not experienced fecundity problems, who mar-

9. Even though parents love their children, they may wish they didn't have them.	.451
*10. Women who don't want at least one child are selfish.	.630
*11. For a woman, marriage should be more important than a career.	.553
*12. It is a woman's moral duty to give her husband at least one child.	.649
*13. Giving birth to a child is a wonderful thing that all women should experience.	.573
*14. A woman can best realize her interests through her husband.	.536
*15. Women were born with the drive to have children.	.388

Starred items were recoded so that a high score (agreement) indicated a traditional sex role orientation.

⁸ Parameter estimates were obtained using ordinary least squares in two steps with corrections made for the inaccuracies in the calculation of standard errors of the metric coefficients and the standardized coefficients for nonrecursive sources.

Table 1. Results of Two-Stage Least-Squares Estimation of Coefficients in Model 1 (N = 703)

Dependent Variable	Predictor Variable	Standardized Coefficient	Unstandardized Coefficient	Standard Error
Labor Force Participation ($R^2 = .10$) †	FERTILITY	-0.403**	-1.200**	0.568
	WORKMO	-0.022	-0.158	0.336
	YRSWORKBM	0.169**	0.249**	0.108
	EDUCMO	-0.018	-0.040	0.110
	DECMARK	0.020	0.169	0.397
	EDUCR	0.145***	0.444***	0.171
	FAMSES	-0.046	-0.037	0.043
	YRSMAR	0.513***	0.704***	0.150
	SPINC	-0.133***	-0.082***	0.028
	RELGOR	—	—	—
	RELGFRHS	—	—	—
	FECUND	—	—	—
	NSIBS	—	—	—
Current Fertility ($R^2 = .23$) †	LFP	0.011	0.004	0.114
	WORKMO	—	—	—
	YRSWORKBM	—	—	—
	EDUCMO	-0.029	-0.022	0.028
	DECMARK	0.023	0.063	0.099
	EDUCR	-0.078	-0.080	0.071
	FAMSES	0.030	0.008	0.012
	YRSMAR	0.386***	0.178***	0.037
	SPINC	0.031	0.006	0.012
	RELGOR	0.183***	0.355***	0.098
	RELGFRHS	0.029	0.025	0.032
	FECUND	0.129***	0.652***	0.170
	NSIBS	-0.032	-0.016	0.022

Correlation of Disturbances (U and V): -0.057

* $p < .10$ ** $p < .05$ *** $p < .01$

† Since the definition of R^2 is problematic for nonrecursive models, they are reported only as a general indicator of the strength of the associations and should not be interpreted in the conventional manner.

ried relatively young, and who had a Catholic upbringing tend to have more children. No other variable even approaches significance. Labor force participation, contrary to expectation, appears to have no causal impact on fertility at age 30.

Model 2

In Model 2, the strongest sources of labor force participation coincide with the findings in Model 1. Education, work experience prior to marriage and marital duration have positive effects. However, in this model fertility has only marginal statistical significance and its influence is weaker than that found in Model 1. This decrease could indicate that part of the negative impact of fertility on work be-

havior is mediated by sex role attitudes (through the coefficient from FERTILITY to SEXROLE and the coefficient from SEXROLE to LFP). But this mediation accounts for only a small portion of the reduction in the size of the coefficient ($-.138 \times .305 = -.042$). Therefore it is likely that the decrease in the FERTILITY to LFP effect is the result of the more accurate specification of Model 2. The correlated errors are no longer being absorbed into the fertility to work coefficient. Finally, there is an effect of sex role attitudes on labor force participation which, while not statistically different from zero, is large enough to be suggestive.

Only three variables significantly influence fertility in Model 2: fecundity, religious preference in high school and mari-

Table 2. Results of Two-Stage Least-Squares Estimation of Coefficients in Model 2 (N = 703)

Dependent Variable	Predictor Variable	Standardized Coefficient	Unstandardized Coefficient	Standard Error
Labor Force Participation ($R^2 = .10$) †	FERTILITY	-0.273*	-0.841*	0.501
	SEXROLE	0.305	0.201	0.134
	WORKMO	0.000	0.002	0.289
	YRWORKBM	0.214***	0.315***	0.102
	EDUCMO	—	—	—
	DECMARK	—	—	—
	EDUCR	0.248***	0.762***	0.258
	FAMES	-0.050	-0.040	0.033
	YRSMAR	0.469***	0.644***	0.127
	SPINC	-0.133***	-0.082***	0.024
	RELGOR	—	—	—
	RELGFRHS	—	—	—
	FECUND	—	—	—
	NSIBS	—	—	—
Current Fertility ($R^2 = .23$) †	LFP	0.090	0.030	0.140
	SEXROLE	0.026	0.006	0.056
	WORKMO	—	—	—
	YRWORKBM	—	—	—
	EDUCMO	—	—	—
	DECMARK	—	—	—
	EDUCR	-0.086	-0.088	0.078
	FAMES	0.019	0.005	0.012
	YRSMAR	0.366***	0.169***	0.051
	SPINC	0.040	0.008	0.014
	RELGOR	0.198***	0.384***	0.134
	RELGFRHS	0.022	0.019	0.047
	FECUND	0.129***	0.654***	0.166
	NSIBS	-0.042	-0.021	0.026
Sex Role Attitudes ($R^2 = .20$) †	LFP	-0.057	-0.087	0.400
	FERTILITY	-0.138	-0.623	1.173
	WORKMO	-0.018	-0.196	0.414
	YRWORKBM	-0.146**	-0.327**	0.161
	EDUCMO	-0.072	-0.242	0.132
	DECMARK	-0.103***	-1.296***	0.479
	EDUCR	-0.317***	-1.476***	0.278
	FAMES	-0.026	-0.032	0.055
	YRSMAR	0.042	0.089	0.346
	SPINC	—	—	—
	RELGOR	-0.059	-0.521	0.546
	RELGFRHS	0.123***	0.490***	0.150
	FECUND	—	—	—
	NSIBS	0.035	0.079	0.095

Correlation of Disturbances (U, V and W):

	U	V	W
U (LFP)	—	—	—
V (FERTILITY)	-0.238	—	—
W (SEXROLE)	-0.302	0.132	—

* $p < .10$ ** $p < .05$ *** $p < .01$

† Since the definition of R^2 is problematic for nonrecursive models, they are reported only as a general indicator of the strength of the associations and should not be interpreted in the conventional manner.

tal duration. Neither labor force participation nor sex role attitudes appears to have any effect on the number of children these women have at age 30.⁹

Model 2 adds a third dependent variable, sex role attitudes. The factors which influence it are R's mother's decision-making power in the family of origin, R's mother's educational attainment, the respondent's own educational attainment, and religious participation in high school. The effect of work experience prior to marriage just misses statistical significance. Frequent religious participation in high school leads to a more traditional sex role orientation, as expected. High educational attainment for mother and respondent and work experience prior to marriage all lead to a more modern set of sex role attitudes.

Neither work nor fertility appear to have much influence on sex role attitudes, although the path from fertility has a standardized coefficient which is comparable in magnitude to the coefficients for some of the significant relationships. The large standard errors for the nonrecursive relationships therefore may be the result of weak instruments.

DISCUSSION

The results of our analyses of fertility and work behaviors differ considerably both from Waite and Stolzenberg's (1976) analyses of plans and from previous speculation in the literature.

It appears that in both Models 1 and 2, background social factors are the most important determinants of the work-fertility relationship. Although the zero-order correlation between the work and fertility variables is -0.458 , much of this relationship is accounted for by the effects of the women's marital duration (for an age cohort, strongly related to age at marriage) and, to a lesser extent, educational attainment on both work and fertility.

The mechanisms which contribute to these effects are related to the timing of life course events (Elder, 1975). Entrance

into marriage at later ages increases the possibility of secondary socialization into extrafamilial adult roles by increasing the probability of educational attainment, work before marriage, and financial independence. Such factors operate both to increase the women's employability and the opportunity costs of not working. Late marriage also leads to lower fertility by increasing the age at which particular parities are attained, and therefore decreasing the probability of each parity progression (Rindfuss and Bumpass, 1976). Indeed, Eckland and Fried (1975), using a panel design to analyze the life histories reported in this data set, found that a large part of the fertility-work correlation appears to be the result of the depressant effect of marriage on both continuing education and employment during the years immediately following high school. This early marriage pattern of limited experience outside the family setting and early fertility is maintained throughout the adult years by carry-over effects. These structural factors therefore appear to set up life course patterns which exert powerful and opposite effects on women's employment careers and childbearing patterns. This creates the spuriously large zero-order relationship which we observe between these two variables during the early years of marriage.

After these structural factors are controlled, however, an effect of fertility on labor force participation remains in both models; there does not appear to be any effect of labor force participation on either current or expected fertility. In other words, the number of children a woman has may influence her work decisions, but her labor force involvement does not alter her fertility behavior. This is the reverse of what Waite and Stolzenberg found when studying *plans* for work and fertility. They found reciprocal causation with a much larger effect from work plans to fertility plans than from fertility expectations to work plans.

There are several factors which might lead to such different patterns of relationships. First, this confirms our earlier assertion that the causal relations between planning variables need not be consistent with the links between behavior variables.

⁹ Additional analyses with expected fertility as the dependent variable show almost exactly the same pattern of results in both models.

Expectations do not entirely determine subsequent actions and behavior may have sources which could not be anticipated at an earlier time (such as unexpected pregnancies, the unavailability of child care, financial setbacks) (Bumpass and Westoff, 1970; Westoff et al., 1957). For example, in a study of the relation between family size preferences of engaged couples and their actual fertility 20 years later, Westoff et al. (1957) found that the correlations ranged from 0.45 (for females using family planning) to 0.12 (for females not using contraceptives). This might indicate that the number of unintended births is an important component of actual behavior measures which is, of course, not included in planning or preference measures.

These considerations are particularly important in view of the fact that the respondents in this analysis are at a different stage in the life course than Waite and Stolzenberg's respondents. Our respondents have been married an average of 11.5 years and 84.6% had two or more children. Clearly, their behavior during the last 15 years indicates any adjustments which they have been forced to make during this period of high time and energy demands from competing role orientations.

Our results indicate that when faced with the actual problem of coordinating work and home life, childbearing had more influence on work behavior than vice versa. Two reasons may be suggested for this pattern. The mother role may have taken precedence for these young, married women because they placed greater importance on the responsibilities and rewards associated with childbearing. After they had one or two children they may have felt constrained to meet normative expectations of childrearing; such demands may have forced them to alter their behavior in other areas such as work outside the home. A closely related explanation emphasizes structural constraints. Some research has indicated that children's schedules are less flexible than work schedules (Darian, 1975). Given that having at least one child is highly valued, the presence of children may have constrained work much more than work con-

strained the number of children. Thus the combination of the structural situation that these women found themselves occupying and the cultural values associated with childbearing and rearing might well explain the stronger effect from fertility to work in the behavioral data. Although expectations about fertility might easily be altered by work plans, the actual inflexibilities of a child's schedule might impose greater barriers to work outside the home than young women (especially those without children) might have supposed.

The inclusion of unplanned births in the behavioral data probably enlarges this effect. Because planned births can be integrated into a working woman's career more readily than unplanned births, the impact of unintended fertility on work will be greater than the impact of planned births. This may explain part of the large coefficient from fertility to work in our data.

In conclusion, our results seem to indicate that the worker and mother roles were to some extent incompatible for these young married women. For this cohort at least, the childbearing and rearing role seems to have taken precedence.¹⁰ Work behavior was influenced by the number of children they had during their twenties; childbearing was not influenced by their work. Additionally, structural factors had a strong impact on both work and fertility behaviors. These relationships may be changing over time (Waite, 1976; Ridley, 1969); therefore complete understanding of the reciprocal influences of work and fertility behaviors will require analyses of other cohorts.

¹⁰ Similar factors may help to explain the weak positive effect of sex role attitudes on work behavior, indicating that women with more traditional role conceptions are more likely to work. This somewhat surprising result may be understood if one considers that these women, who went to high school in the late 1950s and began families in the early 1960s, were very unlikely to have career-potential employment. For them, work was primarily a way to earn extra income for their families rather than a rewarding activity to be pursued for self-enrichment. Perhaps those women who felt most strongly that their proper role was that of family care were more likely to endure the hardships of combining work at a fairly menial job with childcare and housework responsibility when family income without their contribution was low.

REFERENCES

- Alexander, K. L., and B. K. Eckland
 1973 Effects of Education on the Social Mobility of High School Sophomores Fifteen Years Later (1955-1970). Final Report. U.S. Office of Education, Project No. 10202 (00E-4-0037).
 1975 "Contextual effects in the high school attainment process." *American Sociological Review* 40:402-16.
- Barr, A. J., J. H. Goodnight, J. P. Sall, and J. T. Helwig
 1976 A User's Guide to SAS 76. Raleigh: SAS Institute.
- Becker, G. S.
 1976 An Economic Approach to Human Behavior. Chicago: University of Chicago Press.
- Blake, J.
 1970 "Demographic science and the redirection of population policy." Pp. 326-47 in T. Ford and G. DeJong (eds.), *Social Demography*. Englewood Cliffs: Prentice-Hall.
- Bohrnstedt, G. W. and G. Marwell
 1977 "The reliability of products of two random variables." Pp. 254-74 in K. F. Schuessler (ed.), *Sociological Methodology* 1978. San Francisco: Jossey-Bass.
- Boulding, E.
 1976 "Familial constraints on women's work roles." *Signs* 1:95-117.
- Bowen, W. and T. A. Finnegan
 1969 *The Economics of Labor Force Participation*. Princeton: Princeton University Press.
- Bumpass, L. L. and C. F. Westoff
 1970 *The Later Years of Childbearing*. Princeton: Princeton University Press.
- Cain, G. C.
 1966 *Married Women in the Labor Force*. Chicago: University of Chicago Press.
- Darian, J. C.
 1975 "Convenience of work and the job constraint of children." *Demography* 12:245-58.
- Duncan, O. D.
 1975 *Introduction to Structural Equation Models*. New York: Academic.
- Easterlin, R.
 1973 *The Economics and Sociology of Fertility: A Synthesis*. Philadelphia: University of Pennsylvania Press.
- Eckland, B. K. and E. S. Fried
 1975 "A life-cycle model of the relationship between employment and fertility." Paper presented at the annual meeting of the American Sociological Association, San Francisco.
- Elder, G. H., Jr.
 1975 "Age differentiation and the life course." *Annual Review of Sociology* 1:165-90.
- Fishbein, M. and I. Ajzen
 1974 "Attitudes toward objects as predictors of single and multiple behavioral criteria." *Psychological Review* 81:59-74.
- Germain, A.
 1975 "The status and roles of women as factors in fertility behavior." Paper presented at the annual meeting of the Population Association of America, Washington, D.C.
- Goldberger, A.
 1964 *Econometric Theory*. New York: Wiley.
- Groat, H. T., R. L. Workman, and A. G. Neal
 1976 "Labor force participation and family formation: a study of working mothers." *Demography* 13:115-26.
- Heise, D. R.
 1975 *Causal Analysis*. New York: Wiley-Interscience.
- Hoffman, L. and F. I. Nye
 1974 *Working Mothers*. San Francisco: Jossey-Bass.
- Hout, M.
 1977 "A cautionary note on the use of two-stage least squares." *Sociological Methods and Research* 5:335-45.
- Lowenthal, N. H. and A. S. David
 1972 *Social and Economic Correlates of Family Fertility: An Updated Survey of the Evidence*. RTI Project SU-506-9 under USAID Contract NESA/460. Research Triangle Institute, Raleigh, N.C.
- Mincer, J.
 1963 "Market prices, opportunity costs, and income effects." Pp. 67-82 in C. Crist (ed.), *Measurement in Economics*. Stanford: Stanford University Press.
- Namboodiri, N. K.
 1972 "Some observations on the economic framework for fertility analysis." *Population Studies* 26:185-206.
- Piepmeyer, K. and T. S. Adkins
 1973 "The status of women and fertility." *Journal of Biosocial Science* 5:507-20.
- Pratt, L. V. and P. K. Whelpton
 1958 "Extra-familial activities of wives in relation to interest in and liking for children, fertility planning, and actual and desired family size." Pp. 1245-86 in P. K. Whelpton and C. V. Kiser (eds.), *Social and Psychological Factors Affecting Fertility*. New York: Milbank Memorial Fund.
- Ridley, J. C.
 1959 "Number of children expected in relation to non-familial activities of the wife." *Milbank Memorial Fund* 37:277-96.
 1969 "The changing position of American women: education, labor force participation and fertility." Paper presented at the conference on the Family in Transition, National Institute of Health, Washington, D.C.
- Rindfuss, R. R. and L. Bumpass
 1976 "How old is too old? age and the sociology of fertility." *Family Planning Perspectives* 8:5.
- Rosenberg, H. M.
 1973 "Fertility strategies as intervening determinants of wives' labor force status." Paper presented at the annual meeting of the

- Population Association of America, New Orleans.
- Ryder, N. B. and C. F. Westoff
1971 *Reproduction in the United States 1965*. Princeton: Princeton University Press.
- Sastry, K. R.
1975 "Female work participation and fertility." *Journal of Population Research* 2:16-32.
- Simon, J. L.
1974 *The Effects of Income on Fertility*. Monograph 19. Carolina Population Center, University of North Carolina, Chapel Hill.
- Stolzenberg, R. and L. Waite
1977 "Age, fertility expectations and employment plans." *American Sociological Review* 42:769-83.
- Stycos, J. M. and R. H. Weller
1967 "Female working roles and fertility." *Demography* 4:210-7.
- Sweet, J.
1973 *Women in the Labor Force*. New York: Seminar Press.
- Terry, G.
1974 "A theoretical examination of the relationship between fertility and female employment." Paper presented at the Population Association of America, New York.
- Title, C. R. and R. J. Hill
1967 "Attitude measurement and the prediction of behavior: an evaluation of conditions and measurement techniques." *Sociometry* 30:199-213.
- Turchi, B.
1975 *The Demand for Children: The Economics of Fertility in the U.S.* Cambridge, Ma.: Ballinger.
- Waite, L. J.
1976 "Working wives: 1940-1960." *American Sociological Review* 41:65-80.
- Waite, L. J. and R. M. Stolzenberg
1976 "Intended childbearing and labor force participation of young women: insights from nonrecursive models." *American Sociological Review* 41:235-51.
- Westoff, C. F., E. G. Mishler, and E. L. Kelly
1957 "Preferences in size of family and eventual fertility twenty years after." *American Journal of Sociology* 62:491-7.

A RULE FOR INFERRING INDIVIDUAL-LEVEL RELATIONSHIPS FROM AGGREGATE DATA*

GLENN FIREBAUGH
Vanderbilt University

American Sociological Review 1978, Vol. 43 (August):557-572

Under certain conditions aggregate-level data provide unbiased estimates of individual-level relationships. Here I present these conditions in the form of a single theoretical decision rule: bias is absent when, and only when, the group mean of the independent variable (X) has no effect on Y, with X controlled. This paper introduces this rule, demonstrates it for the general n-variable case, compares it with prior discussions of cross-level inference, and illustrates it with the 1930 census data used by Robinson (1950). The final section discusses the implications of this rule for the converse type of cross-level inference: the use of individual-level data to estimate aggregate-level relationships.

Almost three decades ago Robinson (1950) warned sociologists of the dangers of using aggregate data to study individuals. In his seminal paper, Robinson showed that correlations between vari-

ables at the aggregate level differ from correlations between the same variables at the individual level. From this finding Robinson concluded that researchers should not use aggregate data to study individuals;¹ those who did were said to be guilty of the ecological fallacy (Selvin, 1958).

However, as Hammond (1973:765) noted, sociologists (and other social scientists) have continued to use aggregate data

* This rule was first presented, in a rudimentary form, in a fall 1974 methods seminar taught by Karl Schuessler. In refining and extending the ideas of that paper, I received helpful counsel from Leigh Burstein, John Cardascia, Lee Cronbach, Michael Hannan, and Elton Jackson, as well as from Professor Schuessler. Carol Meyer expertly typed and assembled the final draft. Of course, none of the above bears any onus for shortcomings in the final product.

¹ The terms "individual" and "aggregate" refer to units of analysis; an individual need not be a person.

to make inferences about individuals, because appropriate individual-level data are often unavailable. In criminology, for example, deterrence theory suggests that the individual (not the aggregate) is deterred, yet empirical investigations generally have used aggregate data (for example, Tittle, 1969; Chiricos and Waldo, 1970; Logan, 1972; Ehrlich, 1973; 1975). Inference across levels of aggregation (hereafter called cross-level inference)² also can be illustrated by studies of voting behavior, where the use of areal data for studying hypotheses about individuals has a venerable history (see, for example, Burnham, 1965; 1971; and the critique by Cowart, 1974). Other examples—from history, political science, and economics, as well as from sociology—could be given of the substitution of aggregate-level data for unavailable individual-level data.

Given the unavailability of individual-level data for many areas of interest to social scientists, one is not surprised to find that most methodological discussions since Robinson have sought to modify the strict prohibition against downward cross-level inference. The most important conclusion of these discussions has been that aggregate data do not always yield biased estimates of individual-level unstandardized *regression* coefficients (Goodman, 1953; 1959). This conclusion has two implications: first, there are certain cases where, except for possible loss of efficiency (i.e., the variance of $b_{Y\bar{X}}$ is usually greater than the variance of b_{YX} ; see Hannan and Burstein [1974:382]), downward cross-level inference can be made with impunity; second, if downward cross-level inference is made, regression coefficients should be used instead of correlation coefficients.

After the demonstration that an aggregate-level regression coefficient need not differ from its individual-level counterpart, most studies sought to determine the *conditions* under which the regression coefficients do not differ—i.e.,

the conditions under which cross-level bias is absent. Two approaches are discernible. The first approach could be called the contextual effects approach, since it sees contextual effects as the major source of bias (Hammond, 1973; Przeworski, 1974). The second approach is the structural equations approach or causal models approach (Blalock, 1964: 97–114; Hannan, 1971a; 1971b; Hannan and Burstein, 1974; Burstein, 1974; 1975a; 1975b; 1976; Hannan and Young, 1976); this approach formulates bias in terms of path models and uses econometric techniques to determine the expected value of the parameters.³

In this paper I combine the two approaches by employing contextual effects models in a structural equation framework. Two papers in particular stimulated this project: Hammond (1973) and Hannan and Burstein (1974). Hammond's paper is important in that it suggests the link between contextual effects theory and cross-level inference. Hannan and Burstein's paper is important in that it provides a summary of the issues involved in cross-level inference as well as a compact statement of the logic and conclusions of the structural equations approach. Briefly stated, Hannan and Burstein counsel researchers faced with the question of cross-level inference to consider the effects of the variable by which the data are grouped (schools, states, etc.). They show that, in the bivariate case, aggregate data give unbiased estimates of the individual-level relationships when any of the following is true: (1) the grouping variable (A) is uncorrelated with the dependent variable (Y) with the independent variable (X) controlled; (2) A and

² Cross-level inference can be either *downward* (the ecological fallacy) or *upward* (the individualistic fallacy; Alker, 1969: the use of individual-level data to make inferences about aggregate-level effects). Except where otherwise noted, by cross-level inference I mean downward cross-level inference.

³ This literature summary refers mainly to the sociological literature. The econometrics literature most often assumes that the researcher can choose the method of grouping, and discusses the relative efficiencies of different grouping methods (for example, Johnston, 1972: Chap. 7). Recent discussions in political science (Hanushek et al., 1974; Irwin and Lichtman, 1976) have contended that cross-level bias arises from specification error. While the rule introduced in this paper is consistent with this contention, it identifies the source of bias much more specifically than does the generic term "specification error," and thus I judge it to be more useful to the sociologist.

X are uncorrelated; or (3) the variance of X equals the variance of \bar{X} , where \bar{X} is the group mean of X. Hannan and Young (1976) confirm these findings in a Monte Carlo simulation containing two regressors. Burstein (1975a; 1975b; 1976) applies these findings to other empirical examples, and compares the results to those obtained using an approach suggested by Feige and Watts (1972).

The utility of combining the contextual effects and structural equations approaches is shown in this paper. First, the approach adopted here generates a parsimonious rule—the \bar{X} -rule—for making inferences about individual-level relationships from aggregate data. The \bar{X} -rule links cross-level bias to theory on group effects. Since group effects theory is well-known in sociology (e.g., Blau, 1960), this rule often provides theoretical leverage to the sociologist who must determine whether cross-level inference is legitimate in a particular case. Second, this rule is easily generalizable analytically to the n-variable case. This is important since, as Hannan and Young (1976:2) noted, there are formidable obstacles to analytical investigations of the effects of grouping in regression models containing two or more regressors. Finally, unlike most previous approaches, this approach focuses on the difference between the aggregate-level coefficient and the individual-level coefficient of interest (explained in Section III, below).

Like most previous studies, this paper focuses on the conditions for avoiding bias⁴ where the variables of interest are interval scale (for the nominal scale case, see Duncan and Davis, 1953; Shively, 1969; Iversen, 1973). Section I introduces the \bar{X} -rule; first the bivariate case and then the multivariate case are examined. Section II discusses the theoretical interpretation of the \bar{X} -rule, and its implications for research. Section III compares the approach of this paper with previous

approaches to cross-level inference. Section IV provides an empirical illustration. Finally, Section V discusses the implications of this paper for analyses using only individual-level data, as well as summarizing its implications for analyses using only aggregate-level data.

I. THE \bar{X} -RULE: BIVARIATE CASE AND MULTIVARIATE CASE

Introduction of Basic Ideas

We begin with the simplest case: the relationship (slope) between a dependent variable (Y) and a single independent variable (X) in a population. If data on X and Y are available for all individuals in the population, the unstandardized regression coefficient, β_{YX} , is obtained when Y is regressed on X (here, and throughout this paper, the greek letter " β " is used to refer to population parameters; it does *not* refer to standardized regression coefficients). If, on the other hand, all the individuals in the population are placed into mutually exclusive groups (precincts, for example), and an average (usually the mean) for X and Y is computed for each group, the regression of the dependent variable means on the independent variable means yields $\beta_{\bar{Y}\bar{X}}$.

That $\beta_{\bar{Y}\bar{X}}$ is not necessarily equivalent to β_{YX} is well-known to sociologists; the literature is replete with allusions to the danger of inferring individual-level relationships from aggregate-level relationships. Why $\beta_{\bar{Y}\bar{X}}$ may give a biased estimate of β_{YX} , however, is not as well-known. Indeed, a discrepancy between $\beta_{\bar{Y}\bar{X}}$ and β_{YX} may seem counterintuitive, since (1) this discrepancy is not due to sampling error ($\beta_{\bar{Y}\bar{X}}$ and β_{YX} are both population parameters),⁵ and (2) the variables are based on data from the same source (i.e., \bar{X} is computed from data on X, and \bar{Y} is computed from data on Y).

⁴ The question of efficiency is beyond the scope of this paper. The reader should not assume, however, that efficiency is unimportant; biased but efficient estimators often are preferable to unbiased but inefficient estimators. This paper specifies the conditions for unbiased estimation; future papers will want to attend to the issue of efficient estimation.

⁵ As Cronbach (1976:1.9) noted, statistics texts sometimes give the misimpression that aggregation problems involve the issue of inference from sample to population. This is a dangerous misimpression; the reader should clearly distinguish cross-level bias, which involves discrepancies between *population* parameters, from biases which involve discrepancies between sample statistics and population parameters (see also Duncan et al., 1961:62).

The demystification of cross-level bias begins with the recognition that an aggregate variable often measures a different construct than its namesake at the individual level. Often the aggregate-level variable taps more constructs than the individual-level variable. College education is one example (Cronbach, 1976:1.11):

That an individual is college-educated indicates a good deal about what he would be inclined to purchase or what jobs he would be capable of holding. The aggregate college education in the community not only describes an aggregate market and an aggregate employee pool, it says a good deal about what goods and services probably are well-supplied in the community (pediatricians? art movies? books? brokerage offices? etc.), and a good deal about the kinds of jobs offered.

Race is another example: percent black in a community indicates characteristics which are as relevant to the nonblack members as to the black members—their SES, location (urban vs. rural, South vs. non-South), etc. It is this shift in constructs, as one shifts levels of aggregation, which provides the basis for cross-level bias.

When X and \bar{X} measure different constructs, bias is possible. Consider, for example, the finding that, in the 1968 presidential election, percent black was related positively ($r = +.55$) to the Wallace vote for the Congressional districts in the South (Schoenberger and Segal, 1971). In this case, \bar{X} (percent black in Congressional district) no doubt measures extraracial characteristics of the Congressional district which affected the Wallace vote (proximity to Alabama, extensiveness of busing, etc.), thus giving rise to cross-level bias.

On the other hand, consider a hypothetical situation where *all* blacks voted for candidate C, and *all* nonblacks voted for some other candidate. In this case, $\beta_{YX} = 1$ and $\beta_{\bar{Y}\bar{X}} = 1$ (where \bar{Y} = percent of vote for candidate C, and \bar{X} = percent black); hence, $\beta_{YX} = \beta_{\bar{Y}\bar{X}}$. Even though X and \bar{X} measure different constructs—i.e., \bar{X} measures extraracial district characteristics—no cross-level bias results,

since the extraracial characteristics measured by \bar{X} have no effect on Y , with X controlled ("Y, with X controlled," is hereafter written " $Y \cdot X$ ").

We can now state a rule for making downward cross-level inference in the bivariate case:

Cross-level bias is absent when, and only when, $\beta_2 = 0$ in the structural equation $Y = a + \beta_1 X_1 + \beta_2 \bar{X}_1 + e$ (see Figure 1).

This rule is useful to the researcher whether or not s/he can choose the method of grouping. When the researcher has no choice, the question is: is \bar{X} unrelated to $Y \cdot X$? If the researcher can choose between methods of grouping, the question becomes: is \bar{X} unrelated to $Y \cdot X$ under any of the methods of grouping? (It is important to note that \bar{X} may be related to $Y \cdot X$ under one method of grouping but unrelated to $Y \cdot X$ under another method.)

The theoretical interpretation of the rule will be spelled out below. First, however, I give a more rigorous statistical demonstration of the rule; readers who are interested only in its interpretation and application may wish to go directly to Section II.

Bivariate Case

Cross-level bias is the difference, in a population, between the aggregate-level regression coefficient obtained and the individual-level coefficient of interest. In the bivariate case, then, downward cross-level bias (δ) is formally defined as follows:

$$\delta = \beta_{\bar{Y}\bar{X}} - \beta_{YX}, \quad (1)$$

where $\beta_{\bar{Y}\bar{X}}$ is the regression coefficient in the regression of \bar{Y} on \bar{X} (also called the

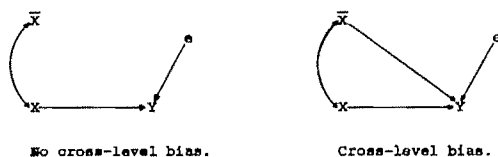


Figure 1. Representation of the Rule for Cross-Level Inference: Bivariate Case.

between-group slope)⁶ and β_{yx} is the regression coefficient in the regression of Y on X within groups (also called the common⁷ within-group slope). (Some writers focus on β_{yx} instead of $\beta_{y\bar{x}}$. However, as we will see in Section III, β_{yx} reflects aggregate-level, as well as individual-level effects.)

Next, consider the following structural model:

$$Y_{ij} = a + \beta_1 X_{ij} + \beta_2 \bar{X}_j + e \quad (2)$$

(i = 1, 2, . . . N; j = 1, 2, . . . M),

where X_{ij} refers to the score on X for the i^{th} person in the j^{th} group, and \bar{X}_j is the group mean of the j^{th} group. This model states that Y is a linear function of X, \bar{X} , and a random disturbance, e, where X and \bar{X} are causal variables (in the theoretical section which follows, we relax the assumption that \bar{X} is causal and consider the case where the effect of \bar{X} is spurious). We make the usual assumptions about e: that it has zero mean, constant variance, is uncorrelated with the independent variables, and that the values of e are mutually uncorrelated. Werts and Linn (1971; also Alwin, 1976) showed that the structural parameters β_1 and β_2 are:

$$\beta_1 = \beta_{yx}; \quad (3)$$

$$\beta_2 = \beta_{\bar{y}\bar{x}} - \beta_{yx}. \quad (4)$$

Equations (3) and (4) state that, in the bivariate case (i.e., in the case of a single X), the individual-level effect of X on Y is β_{yx} , while the aggregate-level effect of X on Y is $\beta_{\bar{y}\bar{x}} - \beta_{yx}$. But $\beta_{\bar{y}\bar{x}} - \beta_{yx}$ also measures downward cross-level bias. Hence, in the bivariate case, downward

cross-level bias is absent when, and only when, X has no aggregate-level effect on Y. Differently stated: the aggregate-level coefficient ($\beta_{\bar{y}\bar{x}}$) provides an unbiased estimate of the individual-level effect of X on Y (β_{yx}) when, and only when, \bar{X} has no effect on Y · X.

This result also can be derived by beginning with the analysis of covariance equation. (Covariance analysis is introduced here since it is easier to generalize to the multivariate case.) The standard covariance equation, with a single covariate, is as follows (see Burke and Schuessler, 1974:165):

$$Y_{ij} = \mu + A_j + \beta(X_{ij} - \bar{X} + e_{ij}) \quad (5)$$

(i = 1, 2, . . . N; j = 1, 2, . . . M),

where μ is common to all cases, A_j is common to all cases in the j^{th} group, X_{ij} is defined as before, \bar{X} is the grand mean of X, and e_{ij} is specific to the i^{th} individual in the j^{th} group. The least-squares solution for the normal equations derived from (5), subject to the constraint that $\sum_{j=1}^M n_j A_j = 0$,

yields the following for the population parameters in (5):

$$\begin{aligned} \mu &= \bar{Y} \text{ (the grand mean of Y);} \\ \beta &= \beta_{yx}; \\ A_j &= \bar{Y}_j - \beta_{yx}(\bar{X}_j - \bar{X}) - \bar{Y}. \end{aligned} \quad (6)$$

Noting that $\bar{Y}_j = a_{\bar{y}\bar{x}} + \beta_{\bar{y}\bar{x}}\bar{X}_j + e_{\bar{y}\bar{x}}$, and substituting (6) into (5), we obtain:

$$\begin{aligned} Y_{ij} &= \bar{Y} + \{(a_{\bar{y}\bar{x}} + \beta_{\bar{y}\bar{x}}\bar{X}_j + e_{\bar{y}\bar{x}}) \\ &\quad - \beta_{yx}(\bar{X}_j - \bar{X}) - \bar{Y}\} \\ &\quad + \beta_{yx}(X_{ij} - \bar{X}) + e_{ij} \quad (7) \\ &= a_{\bar{y}\bar{x}} + \beta_{yx}X_{ij} + (\beta_{\bar{y}\bar{x}} \\ &\quad - \beta_{yx})\bar{X}_j + e. \end{aligned}$$

Hence, whether we begin with equation (2) or with the covariance equation (equation (5)), we conclude that the effect (slope) of \bar{X} on Y · X is $\beta_{\bar{y}\bar{x}} - \beta_{yx}$ (equation (7)).

To summarize: cross-level bias is possible because \bar{X} and X may measure different constructs—an obvious point, but one which has usually been overlooked in the burgeoning literature on cross-level bias. When \bar{X} and X measure different constructs, \bar{X} may affect Y · X. In the bivariate case, an effect of X on Y · X

⁶ The computation of $\beta_{\bar{y}\bar{x}}$ involves weighting by the size of the group:

$$\beta_{\bar{y}\bar{x}} = \frac{\sum_j n_j (\bar{X}_j - \bar{X})(\bar{Y}_j - \bar{Y})}{\sum_j n_j (\bar{X}_j - \bar{X})^2},$$

where \bar{X}_j is the group mean of the j^{th} group, \bar{X} is the grand mean of X, and n_j is the number of individuals in the j^{th} group.

⁷ Here, and throughout this paper, effects are assumed to be linear and additive. In the case of nonadditive relationships, β_{yx} is an inappropriate statistic, since each group should be examined separately (Slatin, 1969).

results in cross-level bias. We will now see that the same principles hold in the multivariate case.

Multivariate Case

Cross-level bias is defined as the difference between the obtained aggregate-level regression coefficient(s) and the individual-level regression coefficient(s) of interest. In the bivariate case, examining cross-level bias involves only one comparison: the comparison of $\beta_{\bar{Y}\bar{X}}$ and β_{YX} . However, if there are n independent variables, n comparisons are involved. With two independent variables (X_1 and X_2), for example, $\beta_{\bar{Y}\bar{X}_1, \bar{X}_2}$ is compared with β_{YX_1, X_2} , and $\beta_{\bar{Y}\bar{X}_2, \bar{X}_1}$ is compared with β_{YX_2, X_1} .

As in the bivariate case, cross-level bias is absent in the multivariate case only when \bar{X} -effects are absent. This condition is represented in Figure 2: \bar{X}_1 has no effect on $Y \cdot X_1, \dots, X_n, \bar{X}_2, \dots, \bar{X}_n$; \bar{X}_2 has no effect on $Y \cdot X_1, \dots, X_n, \bar{X}_1, \bar{X}_3, \bar{X}_n$; etc. The rule for cross-level inference in the multivariate case is as follows:

Cross-level bias is absent when, and only when, $\bar{X}_1, \dots, \bar{X}_n$ have no independent effects in the structural equation

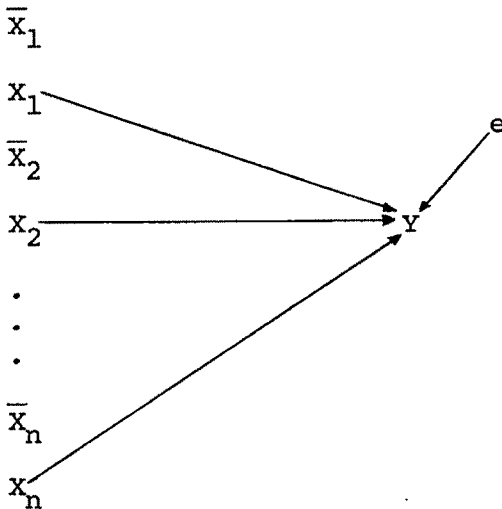


Figure 2. Representation of the Rule for Cross-Level Inference: Multivariate case. (Representation of Zero-Order Correlations between Exogenous Variables Omitted)

$$Y = a + \beta_1 X_1 + \dots + \beta_n X_n + \beta_{n+1} \bar{X}_1 + \dots + \beta_{2n} \bar{X}_n + e.$$

I now give a statistical demonstration of this rule.

To simplify the exposition, we first consider two independent variables (X_1 and X_2); the method then easily generalizes to n independent variables. We begin with the following structural model:

$$Y_{ij} = a + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \beta_3 \bar{X}_{1ij} + \beta_4 \bar{X}_{2ij} + e, \quad (8)$$

where the variables are defined as before (note that an additional subscript is required to distinguish X_1 and X_2), and the assumptions about the error term are the same as in equation (2).⁸ The structural parameters for equation (8) are:

$$\begin{aligned} \beta_1 &= \beta_{YX_1, X_2}; \\ \beta_2 &= \beta_{YX_2, X_1}; \\ \beta_3 &= \beta_{\bar{Y}\bar{X}_1, \bar{X}_2} - \beta_{YX_1, X_2}; \\ \beta_4 &= \beta_{\bar{Y}\bar{X}_2, \bar{X}_1} - \beta_{YX_2, X_1}. \end{aligned} \quad (9)$$

The parameters in (9) can be generated by repeating the procedure used in the bivariate case: (1) begin with the covariance equation; (2) derive the least-squares solution for the parameters (μ , A_j , and β 's); (3) state \bar{Y} as a function of the \bar{X} 's, and substitute the latter for \bar{Y} ; (4) rearrange the terms so that Y is given in terms of the X 's and \bar{X} 's (see the appendix).

Equation (9) shows that $\beta_{\bar{Y}\bar{X}_1, \bar{X}_2} - \beta_{YX_1, X_2}$ gives the effect of \bar{X}_1 on $Y \cdot X_1, X_2, \bar{X}_2$, and $\beta_{\bar{Y}\bar{X}_2, \bar{X}_1} - \beta_{YX_2, X_1}$ gives the effect of \bar{X}_2 on $Y \cdot X_1, X_2, \bar{X}_1$. Hence, as in the bivariate case, the terms attached to the \bar{X} -variables are the cross-level bias terms. Cross-level bias, then, is absent when, and only when, \bar{X}_1 and \bar{X}_2 have no structural effects.

The generalization of this result to n

⁸ Note that mixed models—i.e., models where the \bar{X} -effects involve different variables than the individual-level effects—are possible in the multivariate case if we allow individual-level parameters to be zero. For example, if $\beta_2 = \beta_3 = 0$ in equation (8), the model reduces to $Y = a + \beta_1 X_1 + \beta_4 \bar{X}_2 + e$. Mixed models change the cross-level bias terms; in the current example, the individual-level effect of interest is β_{YX_1, \bar{X}_2} , not β_{YX_1, X_2} . Except where otherwise noted, the discussion throughout assumes non-zero individual-level parameters.

independent variables is straightforward, though notationally cumbersome. We begin with the following structural model:

$$Y_{ij} = a + \beta_1 X_{1ij} + \dots + \beta_n X_{nij} + \beta_{n+1} \bar{X}_{1j} + \dots + \beta_{2n} \bar{X}_{nj} + e, \quad (10)$$

where the variables are defined as before, and the assumptions about the error term and measurement error are the same as in equation (2). The structural parameters for equation (10) are (see the Appendix):

$$\begin{aligned} \beta_1 &= \beta_{yx_1 \cdot x_2, \dots, x_n} \\ &\vdots \\ \beta_n &= \beta_{yx_n \cdot x_1, \dots, x_{n-1}} \\ \beta_{n+1} &= \beta_{y\bar{x}_1 \cdot \bar{x}_2, \dots, \bar{x}_n} - \beta_{yx_1 \cdot x_2, \dots, x_n} \\ &\vdots \\ \beta_{2n} &= \beta_{y\bar{x}_n \cdot \bar{x}_1, \dots, \bar{x}_{n-1}} \\ &\quad - \beta_{yx_n \cdot x_1, \dots, x_{n-1}} \end{aligned} \quad (11)$$

As before, the coefficients of the \bar{X} -terms are the bias terms. Hence, equation (11) demonstrates the rule for cross-level inference in the n -variable case: downward cross-level bias is absent when, and only when, $\bar{X}_1, \dots, \bar{X}_n$ have no structural effects on Y .

II. THE \bar{X} -RULE: THEORETICAL INTERPRETATION AND IMPLICATIONS FOR RESEARCH

When deriving a general rule for cross-level inference, it is necessary to focus on the structural (i.e., "true") relationships between variables. But the researcher of course deals with *observed* relationships; these relationships may be misspecified. Hence, in applying the \bar{X} -rule, the researcher must not only determine if \bar{X} -effects are present, but must also assess whether they are true effects. In this process, the researcher making downward cross-level inference is doubly handicapped: not only must s/he deal with the issue of whether \bar{X} -effects are structural, but s/he must make this assessment without being able to estimate \bar{X} -effects empirically.

This section is intended to help the re-

searcher in applying the \bar{X} -rule in empirical analysis. Applying the \bar{X} -rule involves asking two questions. (1) Are \bar{X} -effects present? (2) If \bar{X} -effects are present, can they be eliminated by respecification of the (aggregate) equation? These issues must be determined by theory. First, then, I discuss the general theoretical underpinnings of \bar{X} -effects.

Group Effects Theory

Sociologists have often argued that groups can (and do) have effects over and beyond those of the attributes of the group members (see, for example, Durkheim, 1897; Merton and Kitt, 1950; Blau, 1957; 1960). Blau (1960:179) expressed this issue as follows:

The individual's orientation undoubtedly influences his behavior; the question is whether the prevalence of social values in a community also exerts social constraints upon patterns of conduct that are independent of the influences exerted by the internalized orientations.

Stouffer et al. (1949), for example, found that inexperienced soldiers in veteran units were less likely to say that they were ready for combat than inexperienced soldiers in inexperienced units—an indication of the influence of the veterans (who generally said they were not ready for combat). Sociologists also have found evidence of group effects in public assistance agencies (Blau, 1960), book discussion groups (Davis et al., 1961) and high schools (Alexander and Eckland, 1975; but see Hauser et al., 1976), to name a few.

Some discussions give the impression that groups are unidimensional and that there is, at most, a single group effect.⁹ But groups can have numerous distinguishable properties even as individuals can have numerous distinguishable properties. Hence I prefer to speak in terms of macroproperties and microproperties. Macroproperties can be divided into two

⁹ This misimpression probably has arisen because the effect of the nominal-scale variable in covariance analysis is often called a group effect. This effect is probably better called the composite group effect, since it measures the total impact of all the group properties affecting Y (Firebaugh, 1977a).

Table 1. Three Cases Where X_1 Would Be Related to $Y \cdot X_1$

Case	Structural Equation	Further Conditions.
Case 1. Spurious Group Effect	$Y = a + \beta_1 X_1 + \beta_2 X_2 + \epsilon$	\bar{X}_1 correlated with X_2
Case 2. Non- \bar{X}_1 Group Effect	$Y = a + \beta_1 X_1 + \beta_2 \bar{X}_2 + \epsilon$ or* $Y = a + \beta_1 X_1 + \beta_2 I + \epsilon$	\bar{X}_1 correlated with \bar{X}_2 \bar{X}_1 correlated with I
Case 3. Emergent \bar{X}_1 -Effect	Disputed ^b	—

* The symbol "I" represents an integral macroproperty.

^b See discussion in text.

classes, according to whether or not they are measured by aggregating microproperties. Macroproperties which are not measured by aggregation—form of government, for example—have been called integral properties (Selvin and Hagstrom, 1963).

Recall the earlier observation that a correlation between \bar{X} and $Y \cdot X$ is possible because \bar{X} can measure more constructs than X .¹⁰ Note three possible reasons for a correlation between \bar{X} and $Y \cdot X$: (1) \bar{X} can reflect microproperties other than X ; (2) \bar{X} can reflect macroproperties other than \bar{X} ; and (3) \bar{X} itself can have a causal effect on Y . We will now see that these cases are central in the discussion of eliminating cross-level bias through respecification.

Eliminating \bar{X} -Effects

Table 1 presents the three cases where \bar{X} would be correlated with $Y \cdot X$. When the \bar{X} -effect is due to uncontrolled individual-level variables (Case 1), cross-level bias can be avoided by respecification of the aggregate equation. On the other hand, the cross-level biases involved in Cases 2 and 3 ordinarily cannot be eliminated by respecification, since they involve causal macroproperties. We now examine each of these cases in more detail.

Case 1. spurious group effect. As Hauser (1970; 1974) has argued convincingly, group effects may be merely individual-level effects in disguise. Spurious \bar{X} -effects occur when \bar{X} is correlated

with uncontrolled causal individual-level variables. Table 1 gives the prototypical form of this case: the causal variables, X_1 and X_2 , are both individual-level variables, and X_2 is correlated with \bar{X}_1 . In this situation, \bar{X}_1 will be correlated with $Y \cdot X_1$ through its correlation with X_2 . This problem can be remedied by respecifying the aggregate equation. Observe that \bar{X}_1 and \bar{X}_2 have no independent effects in a regression of Y on X_1 , X_2 , \bar{X}_1 , and \bar{X}_2 (note the structural equation in Table 1). Now apply the multivariate form of the \bar{X} -rule: if \bar{X}_1 and \bar{X}_2 have no independent effects, then the equation $\bar{Y} = a + \beta_1 \bar{X}_1 + \beta_2 \bar{X}_2 + \epsilon$ provides unbiased estimates of the parameters of the structural equation $Y = a + \beta_1 X_1 + \beta_2 X_2 + \epsilon$.

In short, in the case of spurious group effects, unbiased estimates are possible even when the researcher does not have recourse to individual-level data. How does one know whether group effects are causal or spurious? In the final analysis, the question of whether the effect of a given macroproperty is spurious—like the question of whether the effect of a given microproperty is spurious—must be resolved by theory. As a general rule, causal group effects seem most likely in groups where group members interact and share relevant life experiences; hence, the macroproperties of "natural" groups (neighborhoods, for example) seem more likely to have causal effects than the macroproperties of arbitrarily-created regions (census tracts, for example).

Case 2. non- X_1 group effect. In Cases 1 and 2, \bar{X}_1 is correlated with $Y \cdot X_1$ through its relationship with variables, other than X_1 , which cause Y . In Case 1, these causal variables are microproperties; in Case 2 they are macroprop-

¹⁰ I am grateful to an ASR reviewer for suggesting that I link this observation more directly to group effects theory.

erties.¹¹ Unlike Case 1, then, Case 2 involves causal group effects. The causal macroproperties may be either \bar{X} -properties or integral properties. Table 1 gives the prototypical forms: in the first equation for Case 2, \bar{X}_2 has a causal effect on Y ; in the second equation, I (denoting an integral macroproperty) has a causal effect on Y .¹²

Consider again Cronbach's (1976) example of mean education of communities. To provide a focus for our consideration, suppose that we wish to study opera attendance (Y). Certainly education (X_1) causally affects opera attendance. Other microproperties, such as income, probably also affect opera attendance; let's denote these variables X_2, \dots, X_n . But community macroproperties also could affect opera attendance; for example, community facilities for opera performances (I) likely affects Y . If mean community education (\bar{X}_1), or any other community mean \bar{X}_2 to \bar{X}_n , is correlated with I , the aggregate equation $\bar{Y} = a + \beta_1 \bar{X}_1 + \dots + \beta_n \bar{X}_n + \epsilon$ will give biased estimates of the individual-level effects.

At first blush, the solution to this problem seems straightforward: control the causal macroproperties. However, this approach does not necessarily work. Consider the simplest case: two causal variables, X_1 and \bar{X}_2 (Table 1). Since \bar{X}_1 has no independent effect on $Y \cdot X_1, X_2, \bar{X}_2$, we know that β_1 in the equation $\bar{Y} = a + \beta_1 \bar{X}_1 + \beta_2 \bar{X}_2 + \epsilon$ is an unbiased estimate of $\beta_{YX_1 \cdot X_2}$ (see equation (9), above). However, $\beta_{YX_1 \cdot X_2}$ is not the parameter of interest; we want the effect of X_1 on $Y \cdot \bar{X}_2$, not the effect of X_1 on $Y \cdot X_2$. Unless the effect of X_1 on $Y \cdot \bar{X}_2$ is equivalent to the effect of X_1 on $Y \cdot X_2$, then, the regression of \bar{Y} on \bar{X}_1

and \bar{X}_2 will give a biased estimate of the effect of X_1 .

Case 3. emergent \bar{X} -effect. Consider the following question: since \bar{X} is a sum of X , how can \bar{X} be related to Y once X is controlled? We have already examined two possibilities. \bar{X} could be related to uncontrolled causal microproperties or it could be related to causal macroproperties. Now we consider a third possibility: that \bar{X} itself could give rise to properties which causally affect $Y \cdot X$. That is, Case 3—unlike Cases 1 and 2—involves emergent properties implicit in the aggregation of X .

I am referring in particular to emergent group "atmospheres"; effects of such atmospheres are variously termed "contextual" (Farkas, 1974), "structural" (Blau, 1960), or "compositional" (Davis et al., 1961; Werts and Linn, 1971). We can illustrate such effects with an example from Boudon (1963): assume X = income and Y = voting behavior (conservatism). Boudon suggested that, in France, the mean income of a neighborhood has a positive effect on conservatism, net of individual income. A person living in a high-income neighborhood, then, is expected to be more conservative politically than a person with the same income living in a low-income neighborhood. In such a case, \bar{X} (mean neighborhood income) would be related to $Y \cdot X$.

Social scientists dispute whether \bar{X} should be considered the structural (i.e., true) variable in such a case.¹³ This dispute turns on the question of whether the emergent property generated by \bar{X} should be viewed as a separate variable or as an inherent part of \bar{X} (Cronbach, 1976:1.27: "The properties of what the physicists call a critical mass arise from the aggregate itself, not some 'additional variable'. The whole in this case is more than the sum of the parts").

Consider again the effect of mean neighborhood income on political conservatism. How could such an effect come about? The implicit argument apparently

¹¹ Davis (1966) discussed an effect which is difficult to classify either as micro or macro. This is the "frog pond" effect. In the frog pond effect, group members use some group property as a point of comparison; the individual's position relative to that group standard then affects Y . When frog pond effects involve \bar{X} , \bar{X} will be related to $Y \cdot X$ (Firebaugh, 1977b).

¹² In the case where more than one macroproperty causally affects Y , the correlation between X_1 and $Y \cdot X_1$ is determined by the correlation of X_1 with the composite effect of the macroproperties (Werts and Linn, 1971; Firebaugh, 1977a).

¹³ Indeed, some social scientists question the possibility of emergent properties. I do not care to enter this debate here; the interested reader should consult Hannan (1971a:Chap. 1).

is that a person is influenced by those with whom s/he interacts, and that people in affluent neighborhoods are more likely than those in poor neighborhoods to interact with conservatives.¹⁴ Hence one might be tempted to say that the causal macro-property is the neighborhood's level of conservatism, not its level of income; that is, one would propose that the following equation is the structural equation: $Y = a + \beta_1 X + \beta_2 \bar{Y} + \epsilon$. But, under this specification, the mean conservatism of a neighborhood is caused by the mean income of the neighborhood (this easily can be shown by taking the within-group expected value of Y in the equation just given). Which is the structural variable, then, mean income or mean conservatism?

I am inclined to view mean income as the structural variable; others may disagree. But this is a moot point relative to cross-level bias: whether Y is caused by X and \bar{X} , or caused by X and \bar{Y} , the aggregate-level equations ($\bar{Y} = a + \beta_1 \bar{X} + \beta_2 \bar{X} + \epsilon$ and $\bar{Y} = a + \beta_1 \bar{X} + \beta_2 \bar{Y} + \epsilon$, respectively) obviously cannot be estimated. In short, in the case of emergent \bar{X} -effects, individual-level data are required for unbiased estimation.

This section concludes the exegesis of the \bar{X} -rule. I will use Robinson's (1950) classic example of the relationship between race and illiteracy to illustrate the major principles set forth by the rule. First, however, I address a final issue: why I have chosen to focus on the within-group slope (β_{yx}) instead of the total individual-level slope (β_{YX}).

III. FOCUSING ON β_{yx}

The focus of this paper differs from the focus of prior discussions of cross-level bias. In discussing the effect of X on Y at the individual level, prior discussions in

sociology typically have focused on the following equation:

$$Y_{ij} = a + \beta_{YX} X_{ij} + e. \quad (12)$$

However, the individual-level effect of X on Y is represented by β_{yx} , not β_{YX} ; β_{YX} is a combination of individual-level and aggregate-level effects of X (Duncan et al., 1961:66):

$$\beta_{YX} = \beta_{yx} + E_{XA}^2 (\beta_{\bar{Y}\bar{X}} - \beta_{yx}), \quad (13)$$

where E_{XA}^2 is the correlation between X and A (the grouping variable). Therefore, the researcher should be interested in estimating β_{yx} , not β_{YX} (Cronbach, 1976).

In discussing the rule for cross-level inference, I have found it convenient to refer to the difference between $\beta_{\bar{Y}\bar{X}}$ and β_{yx} as cross-level bias. Many discussions, however, split this difference into two parts ($\beta_{\bar{Y}\bar{X}} - \beta_{YX}$ and $\beta_{YX} - \beta_{yx}$) and focus on the former. Following Hannan and Burstein (1974:387), I will call $\beta_{\bar{Y}\bar{X}} - \beta_{YX}$ "aggregation bias."

From equation (13) one can derive the relationship between cross-level bias (δ) and aggregation bias (θ):

$$\theta/\delta + E_{XA}^2 = 1, E_{XA}^2 \neq 0 \text{ or } 1, \delta \neq 0. \quad (14)$$

From this identity note, first, that cross-level bias is zero when, and only when, aggregation bias is zero, since $\theta = (1 - E_{XA}^2) \delta$, and $0 < E_{XA}^2 < 1$ (cross-level bias is indeterminate when $E_{XA}^2 = 0$ or 1). Hence, in the bivariate case, rules for avoiding cross-level bias apply to aggregation bias, and conversely. Second, note that aggregation bias is always less than cross-level bias. Third, the size of aggregation bias relative to cross-level bias is a function of the correlation between the independent variable and the grouping variable: as E_{XA}^2 increases, the proportion of cross-level bias that is aggregation bias decreases.

An examination of the empirical example given by Hannan and Burstein (1974:Table 2) illustrates these points. Hannan and Burstein used data for 2,676 incoming university freshmen to assess the likely consequences of grouping under various types of grouping variables. Their purpose was to identify those grouping variables which result in the least aggrega-

¹⁴ There is another possibility: perhaps conservatives select housing in wealthier neighborhoods (and liberals select housing in poorer neighborhoods) than expected on the basis of their income. This is an example of what has been termed "grouping by Y " (Blalock, 1964) or "selection by the dependent variable" (Hammond, 1973). The case of selection by the dependent variable, like the case of emergent \bar{X} -effects, results in cross-level bias.

Table 2. Aggregation Bias and Cross-Level Bias under Four Grouping Variables^a

Grouping Variable (A)	E^2_{XA}	Aggregation Bias (θ) ^b	Cross-level Bias (δ)	θ/δ
S.A.T.	.98	-.001	-.037	.02
Father's education	.03	.039	.040	.97
Self-opinion of academic abilities	.28	.060	.084	.72
Achievement test score	.70	.329	1.080	.30

^a See Hannan and Burstein (1974) for a more complete discussion of these data.

^b Variables were standardized before grouping; hence these numbers differ from those reported in Hannan and Burstein's Table 2.

tion bias, and those which result in the most; as expected, they found that grouping by the independent variable (aptitude score on S.A.T. test) was the best method, while grouping by the dependent variable (score on an achievement test) was the worst.

By contrast, our purpose is to compare aggregation bias and cross-level bias. Table 2 compares aggregation bias and cross-level bias for four of Hannan and Burstein's grouping variables.¹⁵ These variables were chosen since they cover the range of values for E^2_{XA} , from a very large correlation (grouping by S.A.T.) to a very small correlation (grouping by father's education). As expected, aggregation bias is less than cross-level bias. Further, θ/δ is inversely related to E^2_{XA} , as expected. A comparison of the two extremes on θ/δ —S.A.T. and father's education—underscores the differences in focusing on aggregation bias instead of cross-level bias: while aggregation bias is quite different under the two methods of grouping, cross-level bias is about the same.

IV. ILLUSTRATION: ROBINSON REVISITED

An empirical illustration should crystallize the ideas presented in this paper. A

reanalysis of Robinson's (1950) classic illustration of the relationship between race and illiteracy seems fitting, since Robinson's paper is the seminal paper for discussions of cross-level bias in sociology. Using 1930 U.S. census data, Robinson computed the correlation between race (black/nonblack) at the individual level and at the regional level. The correlations were .20 and .95, respectively—graphic confirmation of Robinson's contention that aggregate data misestimate individual-level correlations.

Tables 3 and 4 present the data for Robinson's computations. These data are of course nominal scale, but we can still use them to illustrate the principles set forth in this paper. Unlike Robinson, who used measures of correlation, we employ regression coefficients. We first compute β_{YX} . According to these data (Table 3), 16.3% of blacks and 3.1% of the remainder of the population were illiterate in 1930. Letting I =probability of being illiterate, and letting R be a dummy variable for race (1 if black, 0 otherwise), we can write this individual-level relationship between race and illiteracy as follows:

$$I = .031 + .132R. \quad (15)$$

This equation states that the probability of being illiterate is .163 ($= .031 + [.132] [1]$) for a black, and .031 ($= .031 + [.132] [0]$) for a nonblack. The coefficient for R is analogous to a regression coefficient; indeed,

¹⁵ Professor Leigh Burstein generously supplied the information needed to construct Table 2.

Table 3. Race and Illiteracy (000's, Population Ten Years and Older): 1930 U.S.^a

	Black		Nonblack		Total	
	N	%	N	%	N	%
Illiterate	1,514	16.3	2,770	3.1	4,284	4.3
Literate	7,779	83.7	86,661	96.9	94,440	95.7

^a Source: U.S. Census, 1930.

Table 4. Percent Black and Percent Illiterate, by Nine Regions: 1930 U.S.^a

Region	% Black ^b	% Illiterate ^b
1. New England	1.1	3.7
2. Middle Atlantic	4.0	3.5
3. East North Central	3.7	2.1
4. West North Central	2.6	1.4
5. South Atlantic	27.6	8.3
6. East South Central	27.2	9.6
7. West South Central	18.8	7.2
8. Mountain	.9	4.2
9. Pacific	1.1	2.1

^a Source: U.S. Census, 1930.^b Population ten years and older.

regressing I (as a dummy variable: 1 if illiterate, 0 otherwise) on R yields $\beta_{\bar{Y}\bar{X}} = .132$.

Next we compute $\beta_{\bar{Y}\bar{X}}$. Following Robinson, we group by region. To obtain $\beta_{\bar{Y}\bar{X}}$, however, we note that the important regional classification is South/non-South (see Table 4; by "South" I mean regions 5-7). Table 5 presents the aggregate data for race and illiteracy, grouped by region (South/non-South). From Table 5 we can compute the aggregate-level equation analogous to equation (15):

$$\bar{I} = .019 + .26\bar{R}, \quad (16)$$

where \bar{I} = proportion illiterate and \bar{R} = proportion black. Hence, the relationship between proportion black and proportion illiterate by region overstates the zero-order individual-level relationship between race and illiteracy by .128 (= .26 - .132).

From the finding that $\beta_{\bar{Y}\bar{X}} \neq \beta_{YX}$ we can make two inferences. First, we can infer that percent black by regions has an \bar{X} -effect on illiteracy. This \bar{X} -effect is no doubt indirect; percent black of a region is probably correlated with more direct

causes of illiteracy, such as inferior schools. Second, we can infer that the zero-order individual-level relationship between race and illiteracy is not the relationship of interest; i.e., β_{YX} , like $\beta_{\bar{Y}\bar{X}}$, misestimates the individual-level effect of race on illiteracy.

A complete specification of the individual-level and aggregate-level determinants of illiteracy in the U.S. in 1930 is beyond the scope of this paper (but see Hanushek et al., 1974). Nevertheless, the finding that percent black by region has an effect on illiteracy suggests the examination of race and illiteracy, with region controlled (Alker, 1969: 84-5, also suggests this control). Table 6 presents these data; 19.7% of blacks living in the South in 1930 were illiterate, while only 4.6% of blacks living outside the South were illiterate. This suggests that region, not race, was the major determinant of illiteracy. However, illiteracy among nonblacks differed by only 2% between regions (Table 6). Apparently, neither being black, nor living in the South, *in itself* significantly raised the probability of being illiterate. However, being black in the South did. We can see this interaction effect very clearly by translating the above percentages into an equation for illiteracy:

$$I = .026 + .020R + .020S + .131(R*S), \quad (17)$$

where I and R are defined as in equation (15), S is a dummy variable for region (1 for South, 0 otherwise), and $R*S$ is a dummy variable for the interaction between race and region (1 for blacks living in the South, 0 otherwise). The effect of being a black in the South is striking: it raises the probability of being illiterate by .131, net of the additive effects of race and region. Without further data, the interpretation of this interaction effect is ambiguous (discrimination? school segregation? etc.). Note, however, that controlling this interaction reduces the independent effect of race to .02; i.e., net of region and the region/race interaction, being black raises the probability of being illiterate by only .02. In this analysis, in sum, the net individual-level effect of race on illiteracy is .02; the zero-order individual-level relationship between race and illiteracy is

Table 5. Percent Black and Percent Illiterate, South/Non-South: 1930 U.S.^a

Region	Percent Black ^b	Percent Illiterate ^b
South	24.7	8.3
Non-South	3.0	2.7

^a Source: U.S. Census, 1930.^b Population ten years and older.

Table 6. Race and Illiteracy, South/Non-South (000's, Population Ten Years and Older): 1930 U.S.*

	South					
	Black		Nonblack		Total	
	N	%	N	%	N	%
Illiterate	1,416	19.7	1,001	4.6	2,417	8.3
Literate	5,779	80.3	20,972	95.4	26,751	91.7

	Non-South					
	N	%	N	%	N	%
Illiterate	96	4.6	1,771	2.6	1,867	2.7
Literate	2,003	95.4	65,685	97.4	67,688	97.3

*Source: U.S. Census, 1930.

.132; and the aggregate-level relationship between race and illiteracy is .26. As expected, then, β_{yx} misestimates the net individual-level effect of X on Y when $\beta_{\bar{y}\bar{x}}$ misestimates β_{yx} .

V. IMPLICATIONS AND CONCLUSION

Implications

The major conclusion of this paper is that downward cross-level inference can be made without bias when, and only when, \bar{X} -effects are absent. This conclusion has different implications for upward cross-level inference than it does for downward cross-level inference. We now consider its implications for upward cross-level inference; i.e., for the case where the researcher makes inferences about aggregate-level effects from individual-level data.

The conclusions for downward cross-level inference do not apply in a straightforward way to upward cross-level inference. Consider the bivariate case, for example. In downward inference, at issue is the use of an obtained coefficient ($\beta_{\bar{y}\bar{x}}$) to estimate the coefficient of interest at a lower level of aggregation (β_{yx}); in upward inference, at issue is the use of an obtained coefficient (β_{yx}) to estimate the coefficient of interest at a higher level of aggregation ($\beta_{\bar{y}\bar{x}.x}$). (The coefficient of interest at the aggregate level is $\beta_{\bar{y}\bar{x}.x}$, not $\beta_{\bar{y}\bar{x}}$, since the latter is a function of individual-level as well as aggregate-level effects.) Upward cross-level inference can be made with impunity, then, only when $\beta_{yx} = \beta_{\bar{y}\bar{x}.x}$ (note that, since $\beta_{\bar{y}\bar{x}.x} = \beta_{\bar{y}\bar{x}.x}$, $\beta_{\bar{y}\bar{x}.x} = \beta_{\bar{y}\bar{x}} - \beta_{yx}$; see equations (2) and (4), above). That the rules for

downward and upward cross-level inference are not equivalent can be seen as follows: if $\beta_{\bar{y}\bar{x}} = \beta_{yx}$, then $\beta_{\bar{y}\bar{x}} = \beta_{yx} = \beta_{yx}$ (this is easily seen by examining the formulas for these β 's); if $\beta_{\bar{y}\bar{x}} = \beta_{yx} = \beta_{yx}$, then $\beta_{yx} \neq \beta_{\bar{y}\bar{x}.x}$ ($=\beta_{\bar{y}\bar{x}} - \beta_{yx}$), except in the uninteresting case that $\beta_{\bar{y}\bar{x}} = \beta_{yx} = \beta_{yx} = 0$. In short, it is not true, in general, that upward cross-level inference yields unbiased estimates when downward cross-level inference does (or conversely).

A separate discussion of the conditions under which individual-level data provide unbiased estimates of aggregate-level effects would have little utility, since the researcher with individual-level data can estimate directly the individual-level and aggregate-level effects of X on Y (assuming, of course, that the researcher can group the data as desired: for example, computing \bar{X} for nations requires knowing the nationality of the individuals). The researcher with individual-level data, then, most often can include both X and \bar{X} in the equation; through such a multilevel analysis the individual-level and aggregate-level effects of X can be separated (Alwin, 1976; Cronbach, 1976).

The researcher with purely aggregate-level data has fewer options. As we have seen, if \bar{X} -effects are present, aggregate-level regression coefficients give biased estimates of individual-level X-effects. Hence, the researcher who desires to obtain unbiased estimates of individual-level effects from aggregate data must respecify the equation so that \bar{X} -effects are eliminated. This is possible if the \bar{X} -effects are not structural. Hence, if true group effects are so rare or so small that they almost always can be ignored (see Hauser, 1970;

1974), then Hanushek et al. (1974) are correct in their insistence that the researcher restricted to aggregate data should worry primarily about proper specification; the ecological fallacy is itself a near fallacy. If, on the other hand, group effects cannot be dismissed (see Barton, 1970; Farkas, 1974), unbiased estimates of individual-level relationships often will be unobtainable from aggregate data; in such cases, unbiased estimates can be obtained only from multilevel analysis.¹⁶

Summary and Conclusion

This paper has been directed to the sociologist who faces the question of using aggregate-level data to infer individual-level relationships. In this situation, the crucial question is the following: are the means of the independent variables related to the dependent variable, net of the effects of independent variables? I do not wish to claim too much for my \bar{X} -rule; certainly this paper does not solve all the issues in aggregation. Nevertheless, this approach has advantages. First, it links aggregation problems to group effects theory. This not only provides possible theoretical leverage to the researcher puzzling over the legitimacy of downward cross-level inference in a particular case, but it also may demystify the ecological fallacy for some sociologists. Second, the \bar{X} -rule is easily generalizable to the n -variable case. Finally, this approach focuses explicitly on the difference between the aggregate-level coefficient and the individual-level coefficient of interest, β_{yx}^x .

In conclusion, this paper has dealt with cross-level bias by introducing a general rule for making downward cross-level inference. Of course, the researcher concerned about cross-level bias most often does not have individual-level data and thus cannot determine, empirically,

whether the data conform to the rule. This is a problem with downward cross-level inference, to be sure; however, it does not differ in principle from the specification problem faced in all causal analyses. In regression analysis the researcher always makes assumptions about the data used. The validity of these assumptions most often is determined on theoretical grounds; rarely can all the assumptions be tested empirically. This paper suggests that, in an analysis which uses aggregate data to study individual-level relationships, an additional assumption is needed: that there are no \bar{X} -effects. When this assumption is met, aggregate data can provide unbiased estimates of individual-level effects.

APPENDIX

Derivation of Equation (9)

The standard covariance equation, with two covariates, is as follows:

$$Y_{ij} = \mu + A_j + \beta_1 (X_{1ij} - \bar{X}_1) + \beta_2 (X_{2ij} - \bar{X}_2) + e_{ij}, \quad (1A)$$

where the variables are defined as before. The least-squares solution for the normal equations derived from (1A), subject to the constraint that $\sum_{j=1}^M n_j A_j = 0$, yields the following for the population parameters in (1A):

$$\begin{aligned} \mu &= \bar{Y}, \\ \beta_1 &= \beta_{yx_1 \cdot x_2}, \\ \beta_2 &= \beta_{yx_2 \cdot x_1}, \\ A_j &= \bar{Y}_j - \beta_1 (\bar{X}_{1j} - \bar{X}_1) \\ &\quad - \beta_2 (\bar{X}_{2j} - \bar{X}_2) - \bar{Y}, \end{aligned} \quad (2A)$$

where $\beta_{yx_1 \cdot x_2}$ and $\beta_{yx_2 \cdot x_1}$ are the within-group regression coefficients for X_1 and X_2 , respectively; these coefficients give the individual-level effect of X_1 and X_2 on Y . Noting that $\bar{Y}_j = a \bar{y}_{\bar{x}_1 \cdot \bar{x}_2} + \beta \bar{y}_{\bar{x}_1 \cdot \bar{x}_2} \bar{X}_{1j} + \beta \bar{y}_{\bar{x}_2 \cdot \bar{x}_1} \bar{X}_{2j} + e \bar{y}_{\bar{x}_1 \cdot \bar{x}_2}$, and substituting (2A) into (1A), we obtain:

$$\begin{aligned} Y_{ij} &= \bar{Y} + \{ (a \bar{y}_{\bar{x}_1 \cdot \bar{x}_2} + \beta \bar{y}_{\bar{x}_1 \cdot \bar{x}_2} \bar{X}_{1j} \\ &\quad + \beta \bar{y}_{\bar{x}_2 \cdot \bar{x}_1} \bar{X}_{2j} + e \bar{y}_{\bar{x}_1 \cdot \bar{x}_2}) \\ &\quad - \beta_1 (\bar{X}_{1j} - \bar{X}_1) \\ &\quad - \beta_2 (\bar{X}_{2j} - \bar{X}_2) - \bar{Y} \} \\ &\quad + \beta_1 (X_{1ij} - \bar{X}_1) \\ &\quad + \beta_2 (X_{2ij} - \bar{X}_2) + e_{ij} \\ &= a \bar{y}_{\bar{x}_1 \cdot \bar{x}_2} + \beta_1 X_{1ij} + \beta_2 X_{2ij} \\ &\quad + (\beta \bar{y}_{\bar{x}_1 \cdot \bar{x}_2} - \beta_1) \bar{X}_{1j} \\ &\quad + (\beta \bar{y}_{\bar{x}_2 \cdot \bar{x}_1} - \beta_2) \bar{X}_{2j} + e. \end{aligned} \quad (3A)$$

¹⁶ Fortunately, multilevel analysis does not require individual-level data for all groups. If individual-level data are available for randomly selected groups, estimates can be obtained for β_{yx} and β_{yx}^x . Indeed, such a strategy often may be necessary since, in many cases, collecting individual-level data for all groups is prohibitively costly.

Derivation of Equation (11)

The demonstration of (11) is a straightforward extension of the previous case; I indicate here only the key equations. Beginning with the covariance equation with n covariates, we find that the least-squares solution for A_j is:

$$A_j = \bar{Y}_j - \beta_{yx_1x_2} \dots x_n (\bar{X}_{1j} - \bar{X}_1) - \dots - \beta_{yx_nx_1} \dots x_{n-1} (\bar{X}_{nj} - \bar{X}_n) - \bar{Y}. \quad (4A)$$

$$\text{Noting that } \bar{Y}_j = a\bar{y}\bar{x}_1 \dots \bar{x}_n + \beta\bar{y}\bar{x}_1\bar{x}_2 \dots \bar{x}_n\bar{X}_{1j} + \dots + \beta\bar{y}\bar{x}_n\bar{x}_1 \dots \bar{x}_{n-1}\bar{X}_{nj} + e\bar{y}\bar{x}_1 \dots \bar{x}_n,$$

and substituting (4A) into the covariance equation for n covariates, we obtain:

$$\begin{aligned} Y_{1j} &= \bar{Y} + \{ (a\bar{y}\bar{x}_1 \dots \bar{x}_n + \beta\bar{y}\bar{x}_1\bar{x}_2 \dots \bar{x}_n\bar{X}_{1j} + \dots + \beta\bar{y}\bar{x}_n\bar{x}_1 \dots \bar{x}_{n-1}\bar{X}_{nj} + e\bar{y}\bar{x}_1 \dots \bar{x}_n) \\ &\quad - \beta_{yx_1x_2} \dots x_n (\bar{X}_{1j} - \bar{X}_1) - \dots - \beta_{yx_nx_1} \dots x_{n-1} (\bar{X}_{nj} - \bar{X}_n) - \bar{Y} \} \\ &\quad + \beta_{yx_1x_2} \dots x_n (X_{1j} - \bar{X}_1) + \dots + \beta_{yx_nx_1} \dots x_{n-1} (X_{nj} - \bar{X}_n) + e_{1j} \\ &= a\bar{y}\bar{x}_1 \dots \bar{x}_n + \beta_{yx_1x_2} \dots x_n X_{1j} + \dots + \beta_{yx_nx_1} \dots x_{n-1} X_{nj} \\ &\quad + (\beta\bar{y}\bar{x}_1\bar{x}_2 \dots \bar{x}_n - \beta_{yx_1x_2} \dots x_n)\bar{X}_{1j} + \dots + (\beta\bar{y}\bar{x}_n\bar{x}_1 \dots \bar{x}_{n-1} - \beta_{yx_nx_1} \dots x_{n-1})\bar{X}_{nj} + e. \end{aligned} \quad (5A)$$

REFERENCES

- Alexander, Karl and Bruce K. Eckland
1975 "Contextual effects in the high school attainment process." *American Sociological Review* 40:402-16.
- Alker, Hayward R., Jr.
1969 "A typology of ecological fallacies." Pp. 69-86 in M. Dogan and S. Rokkan (eds.), *Quantitative Ecological Analysis in the Social Sciences*. Cambridge, Ma.: MIT Press.
- Alwin, Duane F.
1976 "Assessing school effects: some identities." *Sociology of Education* 49:294-303.
- Barton, Allen H.
1970 "Allen Barton comments on Hauser's 'context and conseq.'" *American Journal of Sociology* 76:514-7.
- Blalock, Hubert M., Jr.
1964 *Causal Inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.
- Blau, Peter M.
1957 "Formal organization: dimensions of analysis." *American Journal of Sociology* 63:58-69.
1960 "Structural effects." *American Sociological Review* 25:178-93.
- Boudon, Raymond
1963 "Proprietes individuelles et proprietes collectives: un probleme d'analyse ecologique." *Revue Française de Sociologie* 4:275-99.
- Burke, Peter J. and Karl F. Schuessler
1974 "Alternative approaches to analysis-of-variance tables." Pp. 145-88 in Herbert L. Costner (ed.), *Sociological Methodology*, 1973-74. San Francisco: Jossey-Bass.
- Burnham, Walter Dean
1965 "The changing shape of the American political universe." *American Political Science Review* 59:7-28.
1971 "Communication." *American Political Science Review* 65:1149-52.
- Burstein, Leigh
1974 "Issues concerning inferences from grouped observations." Paper presented at the annual meeting of the American Educational Research Association, Chicago.
1975a "Data aggregation in educational research: applications." Technical Report No. 1, March. Vasquez Associates. Also presented at AERA, Washington, D.C.
1975b *The Use of Data from Groups for Inference about Individuals in Educational Research*. Ph.D. dissertation, Stanford University.
1976 "Assessing differences between grouped and individual-level regression coefficients." Technical Report No. 10, April. Vasquez Associates. Also presented at AERA, San Francisco.
- Chiricos, Theodore G. and Gordon P. Waldo
1970 "Punishment and crime: an examination of some empirical evidence." *Social Problems* 18:200-17.
- Cowart, Andrew T.
1974 "A cautionary note on aggregate indicators of split ticket voting." *Political Methodology* 1:109-30.
- Cronbach, Lee J.
1976 *Research on Classrooms and Schools: Formulation of Questions, Design, and Analysis*. Occasional paper of the Stanford Evaluation Consortium.
- Davis, James A.
1966 "The campus as a frog pond." *American Journal of Sociology* 72:17-31.
- Davis, James A., Joe L. Spaeth and Carolyn Huson
1961 "A technique for analyzing the effects of group composition." *American Sociological Review* 26:215-25.
- Duncan, O. D., Raymond P. Cuzzort and Beverly Duncan
1961 *Statistical Geography: Problems in Analyzing Areal Data*. Glencoe: Free Press.
- Duncan, O. D. and Beverly Davis
1953 "An alternative to ecological correlation." *American Sociological Review* 18:665-6.
- Durkheim, Emile
[1897] *Suicide*. Glencoe: Free Press.
1951
- Ehrlich, Isaac
1973 "Participation in illegitimate activities: a theoretical and empirical investigation." *Journal of Political Economy* 81:521-65.

- 1975 "The deterrent effect of capital punishment: a question of life and death." *American Economic Review* 65:314-25.
- Farkas, George
1974 "Specification, residuals, and contextual effects." *Sociological Methods and Research* 2:333-63.
- Feige, Edgar L. and Harold W. Watts
1972 "An investigation of the consequences of partial aggregation of micro-economic data." *Econometrica* 40:343-60.
- Firebaugh, Glenn
1977a "Assessing group effects: a comparison of two methods." Unpublished manuscript, Department of Sociology and Anthropology, Vanderbilt University, Nashville.
1977b "Groups as contexts and frog ponds: some neglected considerations." Unpublished manuscript, Department of Sociology and Anthropology, Vanderbilt University, Nashville.
- Goodman, Leo A.
1953 "Ecological regressions and behavior of individuals." *American Sociological Review* 18:663-4.
1959 "Some alternatives to ecological correlation." *American Journal of Sociology* 64:610-25.
- Hammond, John L.
1973 "Two sources of error in ecological correlations." *American Sociological Review* 38:764-77.
- Hannan, Michael T.
1971a *Aggregation and Disaggregation in Sociology*. Lexington: Heath-Lexington.
1971b "Problems of aggregation." Pp. 473-508 in Hubert M. Blalock, Jr. (ed.), *Causal Models in the Social Sciences*. Chicago: Aldine-Atherton.
- Hannan, Michael T. and Leigh Burstein
1974 "Estimation from grouped observations." *American Sociological Review* 39:374-92.
- Hannan, Michael T. and A. A. Young
1976 "Small sample results on estimation from grouped observations." Technical Report No. 24, October. Vasquez Associates.
- Hanushek, Eric A., John E. Jackson and John F. Kain
1974 "Model specification, use of aggregate data, and the ecological correlation fallacy." *Political Methodology* 1:89-107.
- Hauser, Robert M.
1970 "Context and consex: a cautionary tale." *American Journal of Sociology* 75:645-64.
1974 "Contextual analysis revisited." *Sociological Methods and Research* 2:365-75.
- Hauser, Robert M., William H. Sewell and Duane F. Alwin
1976 "High school effects on achievement," Chap. 11 in William H. Sewell, Robert M. Hauser and David L. Featherman (eds.), *Schooling and Achievement in American Society*. New York: Academic Press.
- Irwin, Laura and Allan J. Lichtman
1976 "Across the great divide: inferring individual-level behavior from aggregate data." *Political Methodology* 3:411-39.
- Iversen, Gudmund R.
1973 "Recovering individual data in the presence of group and individual effects." *American Journal of Sociology* 79:420-34.
- Johnston, J.
1972 *Econometric Methods*. 2nd ed. New York: McGraw-Hill.
- Logan, Charles
1972 "General deterrent effects of punishment." *Social Forces* 51:64-73.
- Merton, Robert K. and Alice S. Kitt
1950 "Contributions to the theory of reference group behavior." Pp. 40-105 in Robert K. Merton and Paul F. Lazarsfeld (eds.), *Studies in the Scope and Methodology of "The American Soldier"*. Glencoe: Free Press.
- Przeworski, Adam
1974 "Contextual models of political behavior." *Political Methodology* 1:27-61.
- Robinson, W. S.
1950 "Ecological correlations and the behavior of individuals." *American Sociological Review* 15:351-7.
- Schoenberger, Robert A. and David R. Segal
1971 "The ecology of dissent: the southern Wallace vote in 1968." *Midwest Journal of Political Science* 15:583-6.
- Selvin, Hanan C.
1958 "Durkheim's 'Suicide' and problems of empirical research." *American Journal of Sociology* 63:607-19.
- Selvin, Hanan C. and Warren O. Hagstrom
1963 "The empirical classification of formal groups." *American Sociological Review* 28:399-411.
- Shively, W. Phillips
1969 "'Ecological' inference: the use of aggregate data to study individuals." *American Political Science Review* 63:1183-96.
- Slatin, Gerald T.
1969 "Ecological analysis of delinquency: aggregation effects." *American Sociological Review* 34:894-907.
- Stouffer, S. A., E. A. Suchman, L. C. DeViney, S. A. Star and R. M. Williams, Jr.
1949 *The American Soldier: Adjustment During Army Life*. Princeton: Princeton University Press.
- Tittle, Charles R.
1969 "Crime rates and legal sanctions." *Social Problems* 16:409-23.
- U.S. Bureau of the Census
1930 *Population, Vol. 2: General Report*. Washington, D.C.: U.S. Government Printing Office.
- Werts, Charles E. and Robert L. Linn
1971 "Considerations when making inferences within the analysis of covariance model." *Educational and Psychological Measurement* 31:407-16.

THE VARIABLE ORDER OF EVENTS IN THE LIFE COURSE*

DENNIS P. HOGAN

University of Chicago

American Sociological Review 1978, Vol. 43 (August):573-586

The temporal order in which a man finishes school, starts working, and first marries is an important characteristic of his life course. Ordering patterns are distributed on a scale of the degree of conformity with the normative ordering of events. Major determinants of ordering patterns are identified. While family background is of limited importance for the ordering of events in the life cycle, the manner in which a man spends the years of late adolescence and early adulthood is of critical relevance. College attendance delays marriage, but not by a sufficient amount of time to prevent substantial numbers of men from marrying prior to completing their schooling. Military service is a major disruptive factor in the life courses of men, although the effects of service in the peacetime army have been less deleterious since men have some discretion in its timing. The unique histories of birth cohorts that result from the age-specific conjunction of period events is a crucial exogenous factor in the life course of men. Men for whom the ordering of events is deviant experience higher rates of marital disruption than do other men. This supports the hypothesis that the variable ordering of events in the life course is a contingency of some importance in the life cycle.

Introduction: The Passage to Adulthood

The process through which a group of youths undergoes the passage to adulthood and attains the status of full, productive members of a society, replacing those members lost through death, is known as cohort replacement. Because only one birth cohort undergoes period events at a particular age, each cohort is characterized by a unique history that is reflected in its behavior and in the production of social change (Ryder, 1965; Mason et al., 1973). The concept of cohort re-

placement thus has proven useful in understanding both social continuity and social change.

In some societies the process of cohort replacement is marked by rites of passage through which a youth is formally inducted into the adult ranks. Such rites of passage are most common in primitive societies. In the urban-industrial societies the process has been protracted over the period of adolescence.

While the passage to adulthood no longer is marked by a single dramatic rite, the occurrence of certain events generally does indicate the achievement of adult status. Among American males the most important events in the achievement of adult status have included the completion of formal schooling, the achievement of relative economic independence through the beginning of a full-time first job, and the formation of one's nuclear family of procreation through marriage (Panel on Youth, 1974; Winsborough, 1975).

While the transition to adulthood is a relatively diffuse process in American society, it is not an unregulated process. Each cohort faces normative regulations regarding the appropriate timing for each of these events (Elder, 1974). Thus, members of the society ordinarily are in broad agreement regarding the approximate age at which it is appropriate for a man or a

* The research reported here is an outgrowth of a dissertation written at the University of Wisconsin, Madison, under the supervision of Professor David L. Featherman. I gratefully acknowledge his suggestions and assistance, as well as the comments provided by Hal Winsborough, Donald Bogue, Robert Hauser, Amy Tsui, and the anonymous reviewers. This research has been supported by National Science Foundation Grant GI-31604, "Occupational Changes in a Generation-II," and by NIH Training Grant GM01190 from the National Institute of General Medical Services (NIGMS). Computing facilities at the Center for Demography and Ecology were supported by the National Institute of Child Health and Human Development (NICHD) Grant IPOI-HD05876. I also wish to thank the Manuscript Unit of the Community and Family Study Center for their assistance in preparation of this manuscript. Any opinions, findings, conclusions, or recommendations are solely those of the author and do not necessarily reflect the views of any of the agencies supporting this work.

woman to start working, first marry, have a first child, and so on. As Elder (1974:176) notes, the conjunction of timing norms results in a new norm regarding the appropriate sequential ordering of events:

The concept of the normative timetable implies a preferred sequence of related activities or stages in a line of activity, and thus invites research on the determinants and consequences of deviant sequences or disarrangements[.] . . . the variable order of events is a type of contingency in the life course.

While at least some research has been done on the determinants of age at completion of schooling (Davis and Bumpass, 1976), beginning of first job (Ornstein, 1976), and age at first marriage (Carter and Glick, 1970; Hogan, 1976), the variable sequential order of these events is a topic that has suffered nearly total neglect. Research on the socioeconomic life cycle indicates that the usual pattern for an American male is finishing schooling prior to beginning a first job, although some men do not conform to this pattern (Duncan et al., 1972: chap. 2; Featherman and Carter, 1976; Ornstein, 1976: chap. 2). When the usual age at each event is examined for successive cohorts the evidence suggests that the population is characterized by a statistical norm in that the timing of events shows a progression, with each cohort first completing its schooling, then beginning a first job, and finally marrying (Panel on Youth, 1974; Winsborough, 1975).

The Ordering of Life Cycle Events

This paper demonstrates that the pattern of ordering of schooling, first job, and marriage experienced by men is characterized by degree of conformity with an hypothesized normative pattern. The ability of men to order their life course events in a normative fashion is found to vary by the unique history of the cohort into which they were born and, in particular, by the military service and education experiences of that cohort. Finally, a normative ordering of the events leading from adolescence to adulthood is shown to reduce substantially the chances that

first marriage will end in a separation or divorce.

The basic idea underlying this analysis is that the passage of an American boy from adolescence to adulthood occurs optimally in a socially prescribed fashion when he first finishes his formal schooling, next becomes financially independent through employment at a full-time job, and finally forms a family of procreation by marriage. Such a sequence of events is viewed as "natural" in the American social setting with its custom of neolocally-resident nuclear families. The structures of social institutions are designed for compatibility with such "natural" patterns.

An individual therefore will achieve maximum harmony between his own life style and the social context in which he lives when his life cycle events are ordered in a normative fashion. On the other hand, a man who lives through the sequence of these transition events in a disorderly fashion more frequently is in discord with the social setting. There has been no widespread restructuring of the social institutions involved with education, work, or marriage to increase their compatibility with men having nonnormative orders of schooling, first job, and marriage. While social institutions have made some arrangements to allow for the disorderly sequencing of events (for example, part-time employment and evening classes for employed students, construction of low-cost married student housing, etc.), such arrangements tend to be of a remedial nature.

Hypotheses about the Ordering of Life Cycle Events

Winsborough (1975) has described the duration of time taken by successive cohorts to complete their educations and finish their entries into marriage. (The length of time for a cohort to complete its transition to adulthood is defined as the length of time in which one-quarter of the cohort finishes its schooling and three-quarters first marry.) The long-term trend has been toward a narrowing of the duration of transition to adulthood for success-

ive cohorts. This pattern is subject to period events that uniquely affect the birth cohorts—such as the age-period specific impact of military service requirements on each cohort (Winsborough, 1975; 1976).

One reason for this trend has been the tendency for more schooling (completed at a later age) and earlier marriage among successive birth cohorts. The tendency towards a closer spacing of these events within the cohort has been manifest in its members by a closer spacing of events for each individual and also by a greater likelihood of a disorderly sequencing of events by a man. The association between individual ordering proclivities among the members of a birth cohort and the duration of passage to adulthood for that cohort is quite high. (The multiple correlation between duration of transition for a cohort and the percent of cohort members sequencing their transitions in a disorderly fashion is 0.87 for those cohorts born 1911 to 1944.)

Each birth cohort thus has a unique history that will influence the timing of the life cycle events of its members. We hypothesize that college attendance increases the probability of a disorderly sequencing of events in the early life cycle, as does service in the military. Those birth cohorts (1922–1931) with the least discretion in the timing of their military service, who simultaneously experienced period trends toward more college attendance and an earlier age at marriage, are hypothesized to have had the least success in structuring their passage to adulthood in a preferred sequential order. The consequent disharmonies between their life style patterns and social institutions are expected to produce deleterious social consequences—in particular, higher rates of marital instability.

Data: The OCG-II Survey

The data source for this study of the determinants and consequences of the temporal ordering of schooling, first job, and marriage is the 1973 Occupational Changes in a Generation survey (hereafter, OCG-II) which was carried out in conjunction with the March demographic

supplement to the Current Population Survey. In 1973, the eight-page OCG questionnaire was mailed out six months after the March CPS and followed by mail, telephone, and personal callbacks. The respondents, comprising 88% of the target sample, included more than 33,500 men aged 20–65 in the civilian noninstitutional population. Also, blacks and persons of Spanish origin were sampled at about twice the rate of whites, and almost half the black men were interviewed personally. A complete description of the survey is found in Featherman and Hauser (1975).

Since the timing of first job, as it relates to the completion of schooling and age at marriage, is a key component in the definition of temporal ordering patterns, we will describe briefly the way in which this variable is determined.¹ In the OCG-II interview schedule, following a question on the highest grade of school completed and the date of completion, each respondent was asked to “describe the *first, full-time civilian job* you had *after* you completed your highest grade in school” and the date “you *actually* began this job, *even if* you started the job before you completed your highest grade in school.” No respondents currently enrolled in school answered these first job questions. Consequently, it is not possible to report having begun a first job without also indicating the completion of schooling.²

¹ The age at completion of highest grade of regular schooling and age at first marriage involve none of the special definitional problems of first job, since the events of completion of highest grade of regular schooling and first marriage are readily identified by all respondents and can then be dated.

² A man who worked full time at a civilian job and then discontinued that job in order to finish an additional year of schooling correctly reported as his first job the job he held after completely finishing school. It would perhaps have been preferable for our purposes to define such men as beginning a first job prior to finishing school, but the bias resulting from the OCG-II definition is probably slight. The percentage of men in any given birth cohort defined by the OCG-II procedures as beginning work prior to completing school does not differ greatly from that estimated by the OCG-I procedures (Duncan et al., 1972: chap. 8) under which men with discontinuities in schooling presumably did report themselves as beginning work prior to completing school. In addition, since any man who married during the interval

The Relationship of Age at Completion of Schooling and Age at First Marriage

We begin by demonstrating that age at completion of schooling and age at first marriage are related to each other among individuals in a fashion that is indicative of an ordering norm.³ The null hypothesis of no association between age at the end of schooling and age at marriage could be tested using a conventional chi-square test of association for the two-way contingency table including those variables. However, in order to divide the association between the two variables into a component due to a tendency for marriage to follow schooling and a component of association not related to ordering, we use log-linear modelling techniques. (See Goodman, 1970; 1972b; or Bishop, 1975, for discussion of log-linear models.)

A contingency table of age at completion of school by age at marriage is divided into two subtables—one in which marriage occurs prior to leaving school and the other in which marriage occurs at the same age as, or after, the end of formal education. These two subtables are treated as categories of an ordering dimension, producing the three-way matrix on which the models of Table 1, Panel A are based.

The first line of this panel indicates that the null hypothesis of no association between age at leaving school and age at marriage is rejected ($\chi^2_{LR} = 1994.4$ with 304 degrees of freedom). Of the total association between these two variables about one-third is due to the tendency of men to finish school before marrying (Table 1, Panel A, line 3). Another third of the association is due to the fact that such ordering tendencies are related to a man's position at the age of completion of schooling and age at marriage regimes (Table 1, Panel A, line 5). (This simply indicates that men

who finish school at an early age are more likely to marry after school completion than are men who finish school at a relatively advanced age.)

The association between the timing of school completion and marriage is even stronger where the respective age regimes (i.e., the margins for marriage age and age at leaving school) are allowed to vary among ten-year birth cohorts (compare lines 4 and 6 of Table 1, Panel B). Further analysis indicates that age at marriage is not solely a result of being subject to different marriage regimes before and after the completion of school. Rather, the age-specific marriage rates to which a man is subject depend in part on the age at which he finishes his schooling (Table 1, Panel C). This is equivalent to the observation that age-specific marriage rates are higher among men who have completed school than among those still in school (see Winsborough, 1976).

These models, therefore, verify that age at marriage is associated with age at termination of schooling among individual men. As Elder (1974) suggests, there is a proclivity among men to time marriage so as to follow the completion of schooling—a proclivity that exists apart from the tendency to finish school and to marry during certain age ranges. This provides an empirical justification for our hypothesis that the ordering of transition events between adolescence and adulthood can be classified according to their degree of conformity with a socially preferred sequence. The construction of such a scale of temporal ordering is the step in the analysis to which we now turn.

Construction of a Scale of Temporal Ordering

The scale of temporal ordering developed here is intended to capture the variation in the normativeness of the ordering pattern. We hypothesize that the three events marking the passage to adulthood occur in a normative order only when a man first completes school, next starts work, and lastly marries. A single inversion from this normative rank order occurs when a man either begins a job prior to finishing school or marries prior to be-

of educational discontinuity is classified under the procedures defined later as deviant in his ordering regardless of his first job report, the misclassification of men who worked for an extended period of time during a discontinuity in schooling as normatively ordering their events is reduced further.

³ In order to make this initial demonstration of a temporal ordering dimension as conclusive as possible, the difficult to define timing of first job variable is not included.

Table 1. Log-Linear Tests of Temporal Ordering Relationship of Age at Completion of Schooling and Age at First Marriage, U.S. Males Born 1907-1952^a

Model ^b	χ^2_{LR}	df	p	Δ	χ^2_H/χ^2_T
A. [E] AGEXED BY [M] AGEX1M BY [O] ORDERING Matrix					
1. [E] [M]	1,994.37	304	.000	11.52	100.00
2. [E] [M] [O]	1,392.90	303	.000	9.26	69.84
3. A2 vs. A1 (Net [O])	601.47	1	.000	2.26	30.16
4. [EO] [MO]	665.17	268	.000	6.34	33.35
5. A4 vs. A2	727.73	35	.000	2.92	36.49
B. [E] AGEXED BY [M] AGEX1M BY [C] CHT10X Matrix					
1. [E] [M] [C]	6,780.18	1419	.000	21.99	100.00
2. [EC] [MC]	2,977.13	1279	.000	13.82	43.91
3. [EC] [MC] [EM]	957.18	975	>.5	7.21	14.12
4. B3 vs. B2 (Net [EM])	2,019.95	304	.000	6.61	29.79
5. [EM] [C]	4,911.60	1115	.000	18.81	72.44
6. B5 vs. B1 (Gross [EM])	1,868.58	304	.000	3.18	27.56
C. [E] AGEXED BY [D] DIFFAMAE Matrix					
1. [E] [D]	13,990.05	381	.000	32.51	100.00
[E] AGEXED BY [O] ORDERING Matrix					
2. [E] [O]	9,741.57	16	.000	21.87	69.63
3. C1 - C2	4,248.48	365	.000	—	30.37

^a In this, as in all other tables, the sample cases have been weighted to reflect true population proportions. The estimated sample frequencies have been adjusted to reflect departures from a simple random sample. Certain cells in the above table are structural zeroes; the chi-square statistic and its degrees of freedom are calculated accordingly.

^b AGEXED = Age at completion of highest grade of schooling (0-10/11-12/13/14.../25/26/27-28/29-31/32-35/36-65). For the matrices of Panels A and B this variable is categorized as (0-10/11-12/13/14/.../25/26/27-28/29-30/31-39/40-65).

AGEX1M = Age at first marriage (13-17/18/19.../30/31-33/34-39/40-64).

ORDERING (Schooling completed before marriage/schooling after marriage).

DIFFAMAE = Age at marriage - Age at completion of schooling (LE -6/-5/.../0/1/2/.../15/GE 16).

CHT10X = Ten-year cohorts (1907-16/1917-26/1927-36/1937-46/1947-52).

The notation indicates those marginal tables that are fit (i.e., used to predict cell frequencies) under that model. [E] [M] indicates that the AGEXED margin and the AGEX1M margin are fit. [EM] indicates that the AGEXED by AGEX1M is fit.

χ^2_{LR} is the likelihood ratio chi-square statistic.

df are the degrees of freedom.

p is the probability level that the chi-square statistic is due to chance.

Δ is the index of dissimilarity between the observed sample frequencies and the expected frequencies obtained with that model.

χ^2_H/χ^2_T is percent of the baseline (total) chi-square accounted for by the chi-square statistic of that model.

ginning work, but after the completion of schooling. An extreme, atypical ordering pattern occurs whenever a man marries prior to the completion of his education; this produces at least two inversions in the normative order.

In order to be classifiable on this three-category temporal ordering scale a respondent had to answer OCG-II questions on whether he was enrolled in school as a regular full-time student in March, 1973, had ever held a first job, and had ever married. For those events which he had completed the respondent was asked to

provide the month and year of the event. To be classified on this variable the respondent had to provide information on at least the year during which he had completed each event. Using reported age in March, 1973, we computed the age at each life cycle event.⁴ These ages were then used to array the respondents into the various ordering shown in Table 2. In cases where two events occurred at the same age, the ordering closest to the typical,

⁴ Date of birth was unavailable at the time of the analysis, necessitating the use of age.

Table 2. Percent of Five-Year Birth Cohorts in Each Possible Temporal Ordering Category, U.S. Males Born 1907-1952

Temporal ^a Ordering	Birth Cohort									Total
	1907- 1911	1912- 1916	1917- 1921	1922- 1926	1927- 1931	1932- 1936	1937- 1941	1942- 1946	1947- 1952	
EJ	3.1	3.1	3.3	3.3	3.2	4.1	5.2	9.2	23.9	8.1
EJM	53.4	53.7	54.2	50.9	50.0	49.2	50.0	45.1	25.9	45.9
Total Typical	56.5	56.8	57.5	54.2	53.2	53.3	55.2	54.3	49.8	54.0
EM	2.2	2.4	2.6	2.2	2.2	1.6	1.5	1.2	0.8	1.7
JE	0.4	0.8	0.4	0.8	0.6	0.5	1.2	1.4	3.6	1.3
EMJ	3.3	3.8	4.8	5.4	5.5	5.2	5.6	5.6	2.6	4.6
JEM	7.3	7.3	6.5	6.0	6.7	7.0	6.1	5.9	3.8	6.0
Total Atypical, Marriage after School	13.2	14.3	14.3	14.4	15.0	14.3	14.4	14.1	10.8	13.6
ME	0.3	0.2	0.3	0.6	0.3	0.2	0.1	0.3	0.1	0.3
JME	1.7	1.9	2.7	2.7	3.9	4.1	3.6	3.1	1.1	2.7
MEJ	1.7	3.2	5.2	7.6	7.3	8.1	7.8	7.4	3.2	5.7
MJE	1.2	1.6	2.3	4.0	3.8	3.9	3.5	2.1	0.6	2.4
Total Atypical, School after Marriage	4.9	6.9	10.5	14.9	15.3	16.3	15.0	12.9	5.0	11.1
None	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	2.1	0.4
M	0.0	0.0	0.0	0.0	0.1	0.2	0.2	0.3	0.5	0.2
E	0.5	0.3	0.4	0.3	0.4	0.3	0.5	0.8	2.2	0.8
NA	25.0	21.5	17.2	16.2	16.0	15.6	14.7	17.4	29.7	19.9
Total Non- classifiable	25.5	21.8	17.6	16.5	16.5	16.1	15.4	18.6	34.5	21.3
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
N	1884	2213	2544	2728	2633	2516	2672	3334	4684	25207

^a The ordering of the letters indicates the sequence in which the events occurred, where E stands for completion of schooling, J for beginning of first job, and M for first marriage.

presumed normative pattern was assigned. In cases where two of the three events have occurred, the third event is presumed to occur last, with the respondent classified accordingly. Included among the nonclassifiables, in addition to those with missing data or response inconsistencies, are those who have completed none or only one event. Despite the heavy requirements on the data, most respondents can be successfully assigned a rank on the temporal ordering scale.⁵

⁵ Comparisons of the characteristics of these nonclassifiable men with the characteristics of all those able to be classified and in the same birth cohorts indicated no differences between the two groups that are problematic for the models estimated here. However, rather large percentages are nonclassifiable among the youngest cohort. Also, population coverage of that cohort is relatively incomplete because of the restriction of the Current Population Survey Sample to the civilian population.

Background and Early Achievement Differentials in Ordering Pattern

As indicated earlier, we expect substantial differences among birth cohorts in the proportions of men who succeed in sequentially timing the events leading to adulthood in a normative fashion. As shown in Figure 1, the ordering of school completion, first job, and marriage is strongly related to the year in which a man is born. The pattern is generally curvilinear: the oldest and youngest birth cohorts are most likely to have a typical ordering pattern and least likely to marry before finishing school. The differences are substantial—for example, about 79%

Therefore, men aged 20 to 25 are not included in the log-linear structural equation models which follow. Where descriptive statistics are reported in the text for these young men, they should be interpreted with caution.

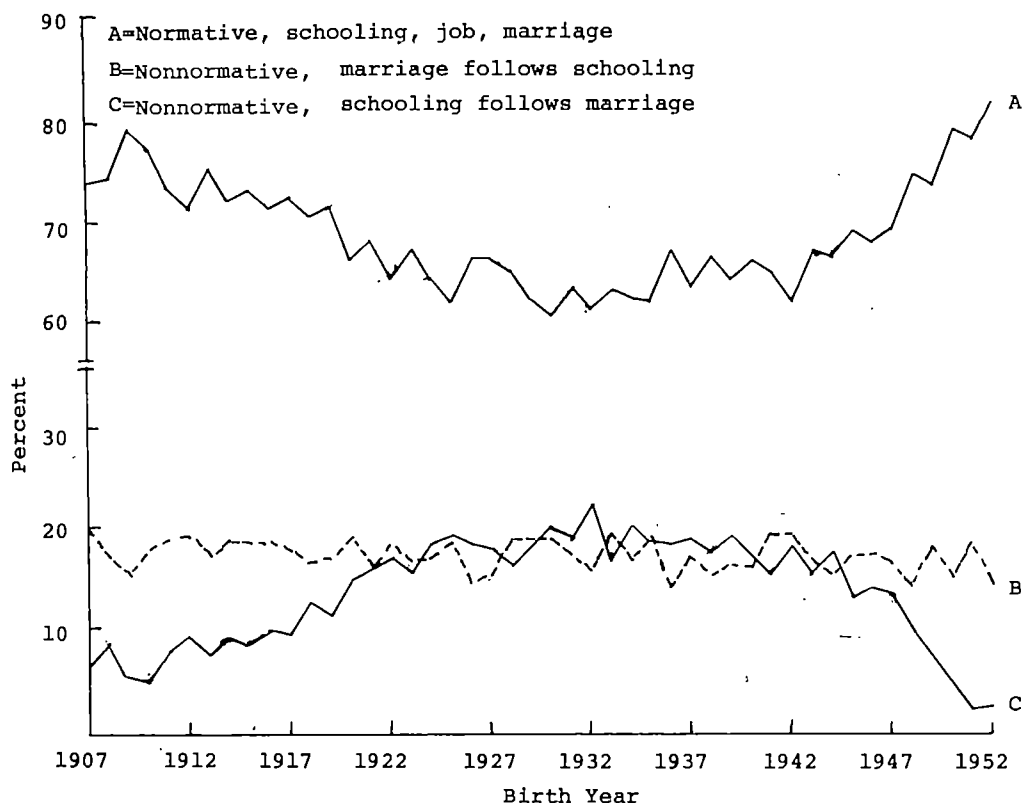


Figure 1. Percent of Birth Cohort in Each of Three Temporal Ordering Categories, U.S. Males Born 1907-1952

of those born in 1909 are normative in ordering compared to less than 61% of the 1930 cohort, with most of the difference accounted for by a greater proclivity among members of the middle cohorts to marry prior to the completion of education. The proportion in the intermediate timing category remains rather constant over all 45 single year birth cohorts.

Higher levels of completed education doubtlessly explain at least part of these cohort differences in ordering pattern. Likewise, military service obligations varied among cohorts and probably account for part of the cohort differences in ordering pattern. In order to decompose the sources of variation in ordering pattern and to measure the net effects of birth cohort, educational attainment, and military service, a log-linear structural equation model is estimated.

The three-category ordering variable is treated as dependent in this model. Birth cohort (classified into five-year groups),

education (categorized into six levels of completed schooling), and military service are included as independent variables in the structural equation model. Because ethnic origin has proved an important differentiating factor in both the socioeconomic life cycle (Blau and Duncan, 1967) and in the process of family formation and dissolution (Carter and Glick, 1970), ethnic ancestry (Anglo, Spanish, black) is included as a control variable in the analysis.

Finally, substantial intercohort upgrading of social origins also might account for some of the differences in ordering pattern among the cohorts. In addition to having indirect effects through educational attainment, such social background factors could produce these differences directly, net of early attainments. To test for this possibility, a series of exploratory canonical correlation and linear regression models that incorporated common measures of social origin (e.g., father's education and

Table 3. Models of Selected Background and Early Achievement Determinants of the Temporal Ordering of Life Course Events, U.S. Males Born 1907-1946

Model ^a	χ^2_{LR}	df	p	Δ	χ^2_B/χ^2_T
A. Baseline Model [EFMCA] [T]	4127.92	1150	.000	17.00	100.00
B. Full Additive Structural Model [EFMCA] [TA] [TC] [TM] [TF] [TE]	751.35	1118	>.5	5.83	18.20
C. Direct Effects Net of All Other Direct Effects					
1. [TA] (ancestry)	10.09	4	.041	0.04	0.24
2. [TC] (birth cohort)	82.18	14	.000	0.51	1.99
3. [TM] (military service)	49.34	2	.000	0.38	1.20
4. [TF] (farm origin)	52.96	2	.000	0.35	1.28
5. [TE] (education)	2954.76	10	.000	10.04	71.58
D. Full Structural Model ^b [EFMCA] [TCM] [TA] [TF] [TE]	701.80	1104	>.5	5.52	17.00
E. Net Effects of Each Three-way Parameter [TCM] (birth cohort-military service)	49.55	14	.000	0.31	1.20

^a See Table 1 for an explanation of the notation used in this table.

^b The full structural model includes all direct effects and all statistically significant ($p < .001$) interactions. The appropriate structural model was selected using reverse stepwise procedures (Goodman, 1971).

occupation, mother's education, family income, number of siblings, etc.) as independent variables were estimated. Due to space limitations we do not report these models here, but the basic finding is that socioeconomic origins and the structure of family of origin only trivially influence a man's ordering pattern. Farm origin is the only family background characteristic that proves to be of relevance for temporal ordering pattern, net of educational attainment and service in the military. Farm origin is consequently included as an independent variable in the log-linear equation model reported here.⁶

The total baseline association of these five independent variables (birth cohort, ethnic ancestry, farm origin, educational attainment, and military service) with the three-category ordering variable that an

appropriate log-linear model must explain is shown in Table 3, Panel A. A full structural equation model which incorporates the direct association of each independent variable with the dependent variable, but does not include joint associations of two or more independent variables with ordering, accounts for 82% of the baseline association (Table 3, Panel B). The component of association due to the direct effect of each independent variable net of the other variables is shown in Panel C. This measure of decrement to the chi-square statistic due to the addition of the direct effect of each variable is analogous to a test for significant increment to explained variance in a linear regression analysis of covariance.

Education, by far the most important variable associated with temporal ordering, accounts for about 72% of the total baseline association. Net of the effect of educational attainment, birth cohort, farm origin, and military service each displays a statistically significant association with temporal ordering. Net of the effects of the other variables, ethnic origin accounts for none of the baseline association of the independent variables with the ordering pattern.

Further insight into the social forces conditioning the ordering of life course events is gained by an examination of the effect parameters which measure the size

⁶ Complete reports on these regression models are found in Hogan (1976). One finding of the canonical analysis worthy of note here is that only one major variate is extracted in a canonical analysis which treats the three-category scale of temporal ordering as the dependent dimension and characteristics of family background and early achievement as independent. Canonical scores extracted to discriminate the categories of the ordering variable on the first variate indicate a normative rank order relationship among the three ordering categories, with the regular (i.e., normative) ordering pattern constituting one pole of the scale and the extreme atypical pattern in which marriage precedes the completion of schooling, the other.

and direction of the impact of each independent variable on the chances of experiencing an ordering pattern of a particular sort. By following the procedures outlined in Goodman (1972a; 1973), the odds parameters (taus) obtained in fitting the structural equation model are translated into the odds parameters (gammas) which are analogous to regression coefficients. These are presented in Table 4. A gamma coefficient equal to 1.00 means that membership in that category of the independent variable does not affect the likelihood of experiencing that type of ordering pat-

tern. A gamma greater than 1.000 indicates a higher likelihood of experiencing that ordering pattern, while a gamma coefficient less than 1.000 signifies a reduced probability of experiencing that ordering pattern among men with the characteristic in question. The larger the departure of the gamma coefficient from 1.000, the greater the impact of that independent variable on the dependent variable. The gammas shown in the "gross effects" column of Table 4 indicate the size of the odds coefficients for each variable when the direct effects of the other independent

Table 4. Structural Models of the Effects* of Selected Background and Early Achievement Variables on the Temporal Ordering of Life Course Events, U.S. Males Born 1907-1946

Independent variables	Gross Effects			Full Additive Structural Model ^d		
	Normative ^b	Intermediate nonnormative	Extreme nonnormative	Normative	Intermediate nonnormative	Extreme nonnormative
(Intercept)	NA ^a	NA	NA	13.418	0.484	.154
Ancestry						
Anglo	0.759	0.683	1.928	1.246	0.908	0.883
Spanish	1.103	1.391	0.652	0.915	1.339	0.816
Black	1.194	1.052	0.795	0.876	0.822	1.388
Birth Cohort						
1942-46	0.864	0.889	1.302	1.056	1.108	0.855
1937-41	0.801	0.756	1.683	0.947	0.906	1.165
1932-36	0.622	0.772	2.081	0.727	0.839	1.639
1927-31	0.652	0.806	1.902	0.721	0.801	1.731
1922-26	0.744	0.803	1.673	0.777	0.717	1.793
1917-21	1.165	1.015	0.846	1.095	0.834	1.095
1912-16	1.689	1.635	0.362	1.340	1.491	0.501
1907-11	2.481	1.793	0.225	1.669	1.668	0.360
Military Service						
No	1.426	0.950	0.738	1.317	0.857	0.886
Yes	0.701	1.052	1.356	0.760	1.167	1.128
Farm Origin						
Nonfarm	0.853	0.629	1.865	1.355	0.797	0.926
Farm	1.173	1.590	0.536	0.738	1.255	1.080
Education						
0-8	8.008	3.699	0.034	8.587	3.203	0.036
9-11	7.077	3.173	0.045	7.316	3.168	0.043
12	2.358	1.864	0.228	2.483	1.865	0.216
13-15	0.297	0.807	4.167	0.292	0.828	4.138
16	0.342	0.231	12.651	0.328	0.242	12.607
17 and up	0.074	0.245	55.476	0.067	0.264	56.597

* These estimated effects are net of the associations among the independent variables. The parameters shown for each ordering pattern refer to the estimated odds of experiencing an ordering of events pattern of that type vs. either other type of ordering.

^b Normative ordering of events occurs when a man first completes school, next begins to work and lastly marries. The intermediate nonnormative pattern occurs when a man begins a job prior to finishing school or marries prior to beginning work, but after the completion of schooling. Extreme nonnormative ordering occurs whenever a man marries prior to the completion of his education.

^c Not shown due to different intercepts for each set of coefficients shown below.

^d This model results in a 12.3% reduction in the conditional uncertainty of type of ordering pattern. The maximum reduction obtainable with this set of variables is 15.1%.

variables are not included in the model. The gammas shown in the "full additive structural model" column are the odds coefficients for each variable, with the direct effects of the other independent variables controlled. (These gross and net gamma coefficients are analogous to unadjusted and adjusted deviations from the mean in a multiple classification analysis.)

The curvilinear relationship of birth cohort to normative ordering pattern described earlier (and displayed in Figure 1) persists with controls for the other independent variables. A normative ordering of events is most characteristic of those men born between 1907 and 1921 and after 1942, with level of educational achievement and service in the military controlled. Also as hypothesized, men who were born between 1922 and 1936 were particularly likely to have continued schooling (or returned to school) following marriage. The intermediate ordering pattern does not follow our expectations, in that it is most common among the oldest and youngest cohorts; it shows a curvilinear pattern similar to that for normative ordering, net of the effects of the other variables.

While Anglos overall are least likely to experience a normative ordering, this is in large part a correlate of the later ages at which Anglos remain in school in attaining higher educations. With level of completed schooling controlled for, the blacks are most likely to display a tendency to marry while in school, whereas Anglo and Spanish men are less likely to do so. (For example, among the Anglos the gross gamma parameter for an extreme non-normative pattern is 1.928, but this is reduced to 0.883 with controls for the other independent variables.) Spanish men are especially likely to begin work prior to finishing school or to marry following schooling but before beginning full-time civilian employment. Overall these patterns result in a higher degree of success among Anglos in the ordering of their passage to adulthood in a socially prescribed fashion.

Men of farm origin are less likely to be in the typical ordering category, net of the effects of all other variables. This is both because they are more likely than men of

nonfarm origin to have a single inversion from the typical pattern and, to a lesser extent, because they are more likely to marry before completing formal education. This curvilinear association of farm origin with normativeness of temporal ordering is due in part to the fact that farm boys are likely to engage in farm work prior to finishing school. Those whose first full-time job was in farming would thus report a date for beginning the job which is prior to the date of the completion of schooling.

As hypothesized, men who served in the military less frequently experience a normative transition to adulthood. They more often have a disorderly sequencing of events, with the intermediate ordering pattern and the extreme nonnormative ordering pattern about equally likely. The total impact of military service is to increase substantially the probability of marriage after schooling, but this is reduced somewhat with controls for birth cohort and educational attainment.

Educational attainment is a major conditioning factor in the determination of a man's ordering pattern. The higher the level of schooling completed by a man, the higher is the probability that he marries before completion of his schooling. The difference is especially pronounced between men with at least some college and men with less than a college education. This simply indicates that there is very little tendency to marry prior to reaching age 18 among any of these men. As anticipated, men attending graduate or professional school during their adult years are especially likely to marry while in school (or prior to resuming interrupted schooling).

The gamma coefficients presented are not intuitively very meaningful for most readers. In order to clarify the nature of the association between educational attainment and temporal ordering patterns, the observed frequencies in that marginal table were examined. Among men with a high school education or less only 5.1% marry prior to finishing school. The comparable figures are 26.5% among men with some college, 35.7% among college graduates, and 57.1% among men with a year or more of graduate or professional

schooling. Conversely, the percentages of each educational group ordering events in a normatively prescribed fashion are, respectively, 76.6, 52.8, 51.5, and 30.2. Thus, while the association of educational attainment with temporal ordering pattern is pronounced, it is far from completely predictable. (In fact, all of the independent variables in the structural equation model together reduce the conditional uncertainty of type of ordering pattern by only 12.3%.)

Our earlier discussion emphasized the unique cohort histories that result in differential discretion among men in the timing of their military service obligations. The unique contribution of postponed schooling among the middle birth cohorts, in conjunction with period tendencies toward earlier marriage, is expected to

make a disorderly sequencing of events especially likely among these men. In other words, military service and higher education both are expected to have enhanced impacts on ordering pattern among men in the middle cohorts. Table 3, Panel E indicates that the only statistically significant interaction is that between birth cohort and military service.

Cohort differentials in temporal ordering vary by whether a man served in the armed forces, net of the effects of all other variables. This interaction is relatively large, accounting for more of the total association in the table than do the direct effects of either military service or farm background. The observed differences in temporal ordering pattern by five-year birth cohorts and military service experience are shown in Figure 2. The cur-

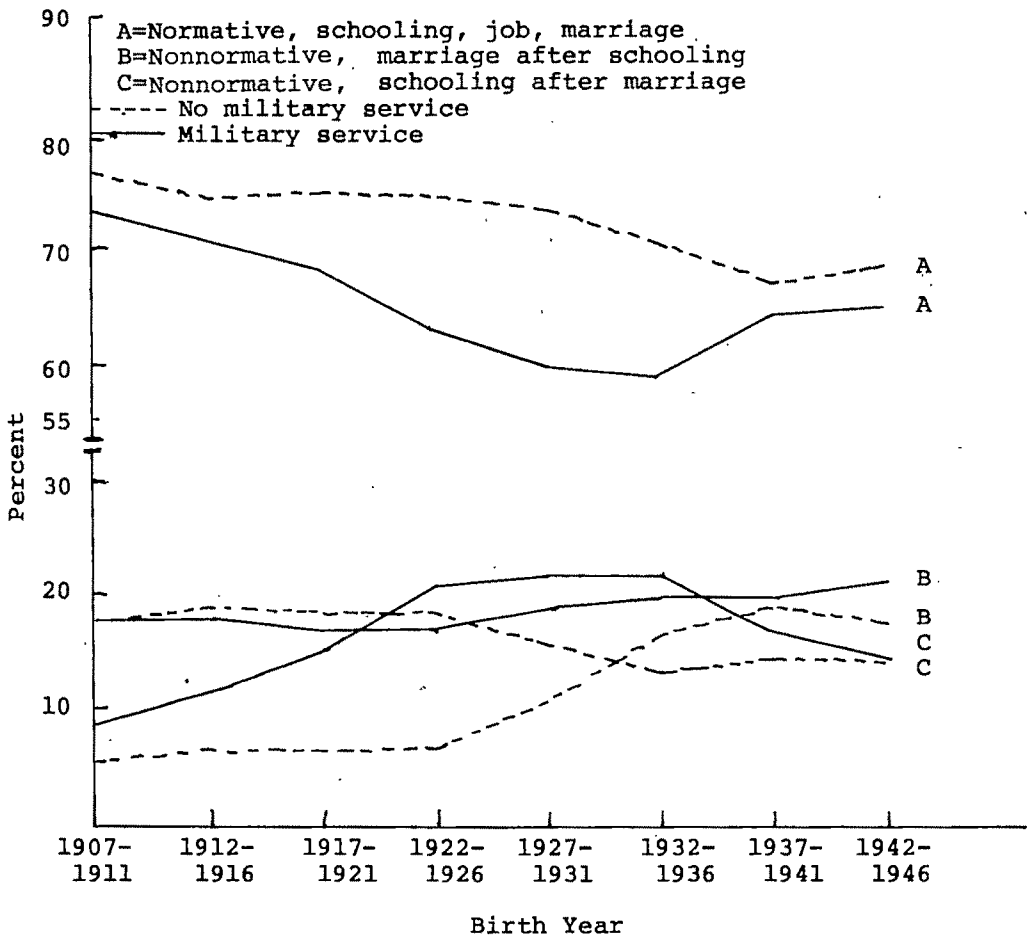


Figure 2. Percent of Five-Year Birth Cohorts in Each of Three Temporal Ordering Categories by Whether They Have Ever Served in the Military

vilinear pattern observed for the total population persists in attenuated form, remaining more marked for veterans than for nonveterans. Within each birth cohort veterans are less likely to be normatively ordered than other men, but the major differences are for those born between 1917 and 1936. Veterans born in these years show a greater tendency to continue school after marriage than do nonveterans. The difference reverses, and remains quite slight, among men born after 1937. From the birth cohorts of 1932 to 1946 veterans were more likely to experience a normative ordering pattern with a single inversion than were men who did not serve in the military. These are the same cohorts among which veterans were becoming relatively less likely to marry before completing their schooling.

Such patterns probably result from the structure of the "peacetime" draft and army, as well as the relative attractiveness of GI Bill benefits and college educations. The men who served in World War II were drawn from cohorts at widely differing ages. That war was a dramatic historical event entailing massive mobilizations of able-bodied men of all ages. Men had little control over the timing of their service. Men born into later cohorts have frequently been called upon to fight in "limited" wars or to serve in the peacetime army. But the foreknowledge of such a military service obligation, with the attendant variety of temporary and permanent deferments, has apparently permitted these later cohorts to adjust the timing of other events in the life cycle to accommodate service in the armed forces, avoiding serious disruption of the sequential life processes.

World War II veterans enjoyed relatively attractive GI Bill benefits and fairly good occupational returns to investments in college training, inducing many married men to complete their interrupted schooling following service in the military. Men born into later cohorts are less likely to marry prior to discharge than were veterans in early cohorts. In recent times, the tendency among married veterans has been to begin work at a civilian job following military service rather than to return to school under relatively meager GI Bill

benefits. (See Hogan, 1976; Winsborough, 1976; Fligstein, 1976, for discussions relevant to these observations.)

The analysis has generally been in accord with our theoretical expectations to this point. We have demonstrated the existence of systematic ordering patterns in the timing of school completion, job beginnings, and first marriage among American men. College attendance increases the probability of a disorderly sequencing of events in the early life cycle, as does service in the military. Those birth cohorts with the least discretion in the timing of their military service frequently did not complete their education and/or began full-time civilian jobs prior to marriage. Period trends toward an earlier age at marriage in the late 1940s and early 1950s enhanced these tendencies toward a disorderly sequencing of life cycle events.

Marital Instability and the Disorderly Sequencing of Events

In the introductory discussion we noted that such disorderly patterning of the transition to adulthood is expected to be consequential for later life chances (Elder, 1974). One widely hypothesized consequence of marriage prior to the completion of schooling and achievement of financial independence is that it increases the likelihood of the first marriage ending in separation or divorce (Bumpass and Sweet, 1975). We now turn to a test of the hypothesis that the variable order of events in the life course should be viewed as a type of contingency with positive consequences (e.g., marital stability) following from a normative ordering pattern and negative consequences (e.g., marital instability) resulting from a deviant ordering of events.

A log-linear structural equation model (not shown here) predicting the likelihood that first marriage is disrupted among ever-married men aged 20 to 65 was estimated. Ethnic ancestry, level of education completed, military service, farm first job, age at marriage, and marriage cohort are included as independent variables, in addition to the three-category ordering of events variable.

Under a log-linear structural equation

model incorporating the direct effects of each of these independent variables on marital instability, the pattern of ordering retains a significant net effect ($\chi^2_{LR} = 14.59$, with 2 degrees of freedom, $p = .000$). The gamma coefficients for the likelihood of marriage ending in a separation or divorce are 0.871 for men with an orderly patterning of events, 1.022 for men with one inversion in the normative ordering pattern, and 1.123 among those men marrying prior to the completion of their schooling. Men who order their events nonnormatively, but who finish school prior to marriage, thus have rates of disruption 17% higher than men who normatively order school, first job, and marriage. Men who finish school after marriage have rates of disruption 29% above those with typical orders. These differentials are exactly the sort expected under the hypothesis that a disorderly sequence of life course events in the transition from school boy to adult male results in negative consequences later in the life cycle.

Summary and Conclusions

This analysis began with a review of the sparse, but suggestive, evidence that the timing of early life cycle events through which American boys pass in their transition to adulthood is governed by an underlying tendency to order those events in a normatively preferred sequence (Elder, 1974). The timing of first marriage among American men proved to be closely associated with age at completion of schooling, with men showing a pronounced tendency to delay marriage until after the completion of schooling.

According to the theoretical framework structuring this analysis, the passage to adulthood is achieved in a socially prescribed manner when a man first finishes his formal schooling, next becomes financially independent through employment at a full-time job, and finally forms his own family of procreation through marriage. The analysis demonstrates that men can be characterized by their degree of conformity with this hypothesized normative ordering pattern.

The ability of men to order their life

course events in a normative fashion varies by the unique history of the birth cohort into which they are born, and in particular by the military service and educational experiences characteristic of their cohort. In general, activities that are time consuming during the period of early adulthood such as the achievement of advanced levels of education (undergraduate and/or graduate and professional training), or service in the military, are disruptive of the normal sequence of life cycle transitions.

Military service and higher education are particularly disruptive of the normal sequence of events among men born during the 1920s. Such men carried the heaviest burden of manpower supply during World War II and thus had little discretion in the timing of their military service. As a consequence, many more of these men married prior to the completion of schooling and/or beginning of work than is the case for men born into different times.

Such a disorderly sequencing of life cycle events is viewed as disrupting the natural harmony between a man's own life style and the institutional structure which provides the social context in which he lives. The consequences of such disharmonies with his social environment persist over a man's life course, substantially increasing the likelihood that his first marriage will end in a separation or divorce.

Birth cohorts born after 1937 have been more successful in reaching adulthood in a normatively prescribed fashion than the preceding cohorts. While college attendance is more common among such men, the discretion they have enjoyed in the timing of their college attendance, and in the timing of their military service obligations (through draft deferrals, early enlistment, or enlistment in the National Guard) permits greater conformity with norms regarding the sequential arrangements of life cycle events. Additionally, period trends toward later age at marriage and more social toleration of premarital sex probably permits such men to postpone marriage until after schooling more readily than could previous birth cohorts.

Continued periodicity in the tendencies of birth cohorts to experience a normative

ordering of events will depend on period trends in educational attainment, age at marriage, and the ease of entry into the labor force. Public policy decisions regarding the future of the voluntary army will continue to play a key role in determining the ability of boys to pass from adolescence to adulthood in a manner that is harmonious with the social environment.

REFERENCES

- Bishop, Yvonne M. M., Stephen E. Fienberg and Paul W. Holland
1975 *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Ma.: MIT.
- Blau, Peter M. and Otis Dudley Duncan
1967 *The American Occupational Structure*. New York: Wiley.
- Bumpass, Larry L. and James A. Sweet
1975 "Background and early marital factors in marital disruption." Paper presented at the annual meeting of the American Sociological Association, San Francisco.
- Carter, Hugh and Paul C. Glick
1970 *Marriage and Divorce: A Social and Economic Study*. Cambridge, Ma.: Harvard University Press.
- Davis, Nancy J. and Larry L. Bumpass
1976 "The continuation of education after marriage among women in the United States: 1970." *Demography* 13:161-74.
- Duncan, Otis Dudley, David L. Featherman and Beverly Duncan
1972 *Socioeconomic Background and Achievement*. New York: Seminar Press.
- Elder, Jr., Glen H.
1974 "Age differentiation and the life course." Pp. 165-90 in Alex Inkeles (ed.), *Annual Review of Sociology*, 1975, Vol. 1.
- Featherman, David L. and T. Michael Carter
1976 "Discontinuities in schooling and the socioeconomic life cycle." Pp. 133-60 in W. H. Sewell, R. M. Hauser, and D. L. Featherman (eds.), *Schooling and Achievement in American Society*. New York: Academic Press.
- Featherman, David L. and Robert M. Hauser
1975 "Design for a replicate study of social mobility in the United States." Pp. 219-51 in K. C. Land and S. Spilerman (eds.), *Social Indicator Models*. New York: Russell Sage.
- Fligstein, Neil D.
1976 *Military Service, the Draft, and the GI Bill: Their Effects on the Lives of American Males, 1940-1973*. Master's thesis, Department of Sociology, University of Wisconsin, Madison.
- Goodman, Leo A.
1970 "The multivariate analysis of qualitative data: interactions among multiple classifications." *Journal of the American Statistical Association* 65:226-56.
- 1971 "The analysis of multidimensional contingency tables: stepwise procedures and direct estimation methods for building models for multiple classifications." *Technometrics* 13:33-61.
- 1972a "A modified multiple regression approach to the analysis of dichotomous variables." *American Sociological Review* 37:28-46.
- 1972b "Some multiplicative models for the analysis of cross-classified data." Pp. 649-96 in L. LeCam (ed.), *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley: University of California Press.
- 1973 "Causal analysis of data from panel studies and other kinds of surveys." *American Journal of Sociology* 78:1135-91.
- Hogan, Dennis P.
1976 *The Passage of American Men from Family of Orientation to Family of Procreation: Patterns, Timing, and Determinants*. Ph.D. dissertation, Department of Sociology, University of Wisconsin, Madison.
- Mason, Karen O., Wm. M. Mason, H. H. Winsborough, and W. K. Poole
1973 "Some methodological issues in cohort analysis of archival data." *American Sociological Review* 38:242-58.
- Ornstein, Michael D.
1976 *Entry into the American Labor Force*. New York: Academic Press.
- Panel on Youth
1974 *Youth: Transition to Adulthood*. Report of the Panel on Youth of the President's Science Advisory Committee (1972). Chicago: University of Chicago Press.
- Ryder, Norman B.
1965 "The cohort as a concept in the study of social change." *American Sociological Review* 30:843-61.
- Winsborough, Hallman H.
1975 "Statistical histories of the life and cycle of birth cohorts: the transition from school boy to adult male." Center for Demography and Ecology Working Paper No. 75-19, University of Wisconsin, Madison.
- 1976 "The transition from school boy to adult: accounting for change in the process." Paper presented at the annual meeting of the Population Association of America, Montreal.

COMMENTS

ON THE MEASUREMENT OF SEGREGATION AS A RANDOM VARIABLE*

(COMMENT ON CORTESE, FALK, AND
COHEN, ASR AUGUST, 1976)

Introduction

Since Duncan and Duncan's (1955a) article, the index of dissimilarity (D) has been used widely as a measure of residential segregation (see Duncan and Duncan, 1955b; Duncan and Lieberson, 1959; Lieberson, 1961; Uyeki, 1964; Taeuber and Taeuber, 1965; Zelder, 1970). However, in spite of its widespread use, researchers have never been comfortable with D as a measure of segregation. Among the many difficulties noted, two frequent objections are that D is affected by differences in the proportion minority (q) in different populations, and by differences in the size (population) of the areal units of analysis (T_1). As a result, the validity of intercity comparisons is often questioned (Duncan and Duncan, 1955a; Taeuber and Taeuber, 1965).

In an effort to correct these and other problems, Cortese et al. (1976) have proposed an alternative measure of residential segregation. However, the measure they propose is inappropriate for several reasons to be discussed below. As a result, the measure does a worse job of measuring segregation empirically than the index it was designed to replace.

Segregation as a Random Variable: Basic Issues and Empirical Results

Cortese et al. (1976) suggest that the inadequacies of D as a measure of segregation stem from the fact that it measures segregation against an arbitrary, a priori expectation of evenness in residential distribution. Instead of evenness, they propose that segregation be measured as departure from an expectation of randomness. To accomplish this, they assume that each census unit (tract, block) represents a random sample taken without replacement

from a city's population. Given the proportion minority in the city and the sizes of the various census units, the randomly expected value of D and its variance may be determined from the hypergeometric probability distribution. The appropriate measure of residential segregation may then be calculated as the standard score of D :

$$Z_D = \frac{D - E(D)}{\sqrt{\text{VAR}(D)}}$$

Cortese et al. maintain that this procedure increases the validity of intercity comparisons by taking into account varying minority proportions and tract sizes. However, taking Z_D as a measure of segregation in no way removes or controls for their effects. $E(D)$ and $\text{VAR}(D)$ are still direct functions of minority proportion and tract size, and to the extent that they vary from city to city, Z_D may also be expected to vary. Extent of departure from either randomness or evenness depends in both cases on q and T_1 . If anything, we expect the value of Z_D to be more sensitive to changes in q and T_1 , since the effect of changes in these factors is to produce independent variation in $E(D)$ and $\text{VAR}(D)$ through the action of the hypergeometric probability function, in addition to any effect on D itself. As for D being based on an arbitrary standard of evenness, ultimately both randomness and evenness are necessarily a priori assumptions, and neither is inherently more correct as an absolute base of comparison.

Another drawback of Z_D as a measure of segregation is that it has no clear intuitive interpretation. Standard deviation units do not mean anything in the context of residential segregation itself. The fact that a particular value of D lies some number of standard deviations above expectation does not imply any clear social meaning. Rather, it is the departure from evenness that bears the social and economic ramifications that make residential segregation important and interesting. As Taeuber and Taeuber (1965:233) point out: "The deviation of residential distribution from evenness, regardless of whether the deviation is larger than anticipated on the basis of random factors, may have distinct consequences for other aspects of the social position of the segregated group."

A basic problem with the procedure proposed by Cortese et al. is that it confuses statistical and substantive significance. A large

* This study was made possible in part by National Science Foundation Grant SOC77-07081, whose support is gratefully acknowledged. The author would also like to thank Julian Wolpert, James Beniger, and Bryan Boulier for their helpful comments and criticisms.

value of Z_D does not necessarily indicate a high level of residential segregation. It is empirically possible for a D of 0.08 to lie several standard deviations above expectation in a large urban population with a high proportion minority, yet it is difficult to see how a D that low could be taken to indicate a meaningful level of segregation. It is the extent of actual unevenness that is substantively important, not how close the unevenness is to what we would expect if settlement were random. The two are really separate questions.

The distinctness of the separate questions underlying the calculation of D and Z_D is readily apparent empirically. As part of a larger research project examining residential location patterns of nonlabor force populations, indices of dissimilarity were calculated between the labor force population and various categories of needy elderly and disabled, people who are permanently out of the labor force. In all, 108 separate values of D and Z_D were calculated for various groups of nonlabor force participants in nine New Jersey counties. A representative set of indices taken from the larger set of 108 is presented in Table 1, showing the residential segregation of disabled males from males who are in the labor force. The areal units of analysis were census tracts, giving a measure of residential segregation for each of the nine counties. D was calculated in the usual manner, and $E(D)$ was computed from equation (11) in Cortese et al. (1976). However, the very large binomial expansions and factorials that must be computed to obtain $VAR(D)$ using the method of Cortese et al. render it completely infeasible for practical application. Therefore an approximation technique developed by Mosteller and Tukey (1968; 1977) was used. This technique—the jackknife—is an empirical method designed to estimate error variability

and approximate confidence limits when standard methods are biased or inapplicable.

If Z_D were indeed a measure of residential segregation, as Cortese et al. suggest, we would expect some correlation when compared to the index of dissimilarity. The important point to note in Table 1 is the complete lack of correspondence between D and Z_D as measures of segregation. For this table, the coefficient of determination between D and Z_D is 0.03, and the rank order coefficient is only 0.05. When all 108 indices are considered, a similar lack of association is found. Interval and rank order coefficients of determination are 0.00 and 0.01, respectively.

Thus the two measures are not at all related, even as ordinal scales of segregation. This is compelling empirical evidence that Z_D is not itself an appropriate measure of residential segregation. In spite of its weaknesses, researchers agree that D does measure segregation in some way, so that if Z_D were in fact a legitimate measure of segregation, at least some correlation with the index of dissimilarity would be expected. To maintain that D and Z_D are both measures of segregation in the face of this evidence would be to suggest that they measure wholly orthogonal dimensions of segregation, a possibility that seems very unlikely. The only realistic conclusion is that Z_D is not a proper measure of residential segregation.

Segregation as a Random Variable: An Alternative View

Cortese et al. proposed Z_D as a measure of segregation rather than a test of the hypothesis that a given pattern of segregation was due to chance. However, for purposes of theoretical explication, it is useful to have some notion of the extent to which a given D may be regarded as a result of a random process. Thus we may alternatively consider Z_D to be a rough test of the null hypothesis that residential location is random with respect to minority group membership. Z_D would thus indicate the degree to which we are able to eliminate a random process model as a plausible theoretical alternative. Having dispensed with chance settlement as an explanation, investigators could then center on D as a measure of residential segregation. The only problem is that the sampling distribution of D is unknown, so that p -values and confidence intervals cannot be determined. However, Z_D may still be useful as a rough guide, and results discussed below suggest that if D is large enough to be of substantive interest, it is typically statistically significant as well.

Table 1. Indices of Dissimilarity, Their Expected Values, Standard Errors, and Standard Scores, Showing the Residential Segregation of Disabled Males from Labor Force Males in Nine New Jersey Counties

County	D	E(D)	SE	Z
Somerset	0.8217	0.0321	0.0615	12.83
Mercer	0.7932	0.0457	0.0523	14.29
Middlesex	0.7754	0.0620	0.0288	24.77
Morris	0.7704	0.0509	0.0487	14.77
Union	0.7575	0.0478	0.0405	17.52
Passaic	0.7035	0.0431	0.0489	13.50
Hudson	0.6982	0.0776	0.0475	13.06
Essex	0.6900	0.1110	0.0403	14.37
Bergen	0.6395	0.1415	0.0224	22.23
$r^2_{DZ} = 0.03$				

Suppose that the areal units of analysis are census tracts. Urban census tracts generally have a population of from 3,000 to 6,000 people, with an average size of around 4,000 (U.S. Bureau of the Census, 1970). When $E(D)$ is computed from equation (11) in Cortese et al., we find that it is near zero for virtually all minority proportions within this size range. Even if q is only 0.01—an atypical circumstance for most applications— $E(D)$ is less than 0.08 at size 3,000 and around 0.05 at size 6,000. When households are considered, census tracts typically range in size from 1,000 to 3,000. Within this size range, all but the lowest minority proportions hover at an $E(D)$ of 0.05 or less. Thus under almost all circumstances using tract data, $E(D)$ is essentially zero.

Z_D is also a function of the standard error of D . Specifically, the smaller the value of SE_D , the larger the value of Z_D . However, because of the infeasibility of computation, we cannot generate theoretical expectations for $VAR(D)$ based on hypergeometric probabilities. Therefore we must rely upon empirical estimates obtained by the jackknife technique. Referring back to Table 1, note that the estimates of standard error are of a very small order of magnitude. Indeed, for all 108 computations, the median value was 0.019, with values ranging from 0.007 to 0.089. Ninety percent of the values were less than 0.04. These values are conservative in that they are based on tract sizes and minority proportions much smaller than would typically be encountered in studies of racial segregation. They are thus computed under conditions that would tend to maximize error variation. Thus the small size of the estimates computed here suggests that a very small order of magnitude is typical for the standard error of D .

We have thus shown that under usual conditions of analysis at the tract level, the expected value of D is near zero and the order of magnitude of its standard error very small. This means that, typically, any index of dissimilarity that is of reasonable substantive interest may be considered to be statistically significant as well. Only under very atypical conditions would we expect a substantively interesting D not to be several standard errors above its expected value.

This proposition is readily born out empirically, as can be seen by glancing at Table 1. The lowest value of Z_D in the table shows D to lie more than 12 standard errors above expectation. Similarly, when all 108 measures are considered, Z -scores are so universally large as to cast serious doubt on the plausibility of a random process explanation. For all cases, the median Z -score was 12.83, with values ranging

from 3.66 to 27.63. Ninety percent of the values were above $Z_D = 8.00$. Thus even though a direct test of statistical significance cannot be computed for lack of knowledge of the sampling distribution of D , we can typically reject a random process explanation with some confidence.

The groups used to compute the above measures provide an excellent test of the degree to which we are able to attribute statistical significance to substantively interesting results because: (1) the total population of each census tract is small, averaging from 1,000 to 2,000 people, much smaller than the 3,000 to 6,000 typically involved in the calculation of indices of racial segregation; and (2) the proportion minority is very small, ranging from 0.01 to 0.09, depending on the particular subgroup under consideration. It is precisely under these conditions that we expect $E(D)$ and $VAR(D)$ to be large, in a sense maximizing the possibility that a large D is a random result.

The above work is thus consistent in suggesting that we are justified in assuming that almost any substantively interesting D is not a result of a random process alone. The ability to assume that a given index of dissimilarity is not a random result is important because the calculation of Z_D is costly and time-consuming, even using high-speed computers. The above discussion suggests that this calculation is unnecessary in virtually all situations where D is calculated from tract data.

Summary and Conclusions

This paper has presented logical and empirical evidence to show that Z_D is an inappropriate measure of residential segregation. Logically, Z_D in no way controls for the effects of varying tract sizes and minority proportions. Moreover, it has no clear theoretical or intuitive meaning with respect to residential segregation. Indeed, the use of Z_D as a measure of segregation confuses the concepts of statistical and substantive significance. The inappropriateness of Z_D as a measure of segregation is born out empirically by the lack of any kind of association between Z_D and D .

On the other hand, Z_D may be useful as a rough guide in assessing the degree to which observed patterns of unevenness may be a result of random processes. But even in this limited role as a quasi-statistical test, the calculation of Z_D is usually unnecessary. Results suggest that except for very atypical situations, if an index of dissimilarity is large enough to be of substantive interest, the expected value and variance will be of such a low order of mag-

nitude that its statistical significance is naturally assured.

Douglas S. Massey
Princeton University

REFERENCES

- Cortese, Charles F., R. Frank Falk, and Jack K. Cohen
1976 "Further considerations on the methodological analysis of segregation indices." *American Sociological Review* 41:630-7.
- Duncan, Otis D., and Beverly Duncan
1955a "A methodological analysis of segregation indexes." *American Sociological Review* 20:210-7.
1955b "Residential distribution and occupational stratification." *American Journal of Sociology* 60:493-503.
- Duncan, Otis D., and S. Lieberman
1959 "Ethnic segregation and assimilation." *American Journal of Sociology* 64:364-74.
- Lieberman, S.
1961 "The impact of residential segregation on ethnic assimilation." *Social Forces* 40:52-7.
- Mosteller, Frederick, and John W. Tukey
1968 "Data analysis, including statistics." Pp. 80-203 in G. Lindzey and E. Aronson (eds.), *The Handbook of Social Psychology*. Reading: Addison-Wesley.
- 1977 *Data Analysis and Regression: A Second Course in Statistics*. Reading: Addison-Wesley.
- Taeuber, Karl E., and Alma F. Taeuber
1965 *Negroes in Cities*. Chicago: Aldine.
- U.S. Bureau of the Census
1970 *Census Users' Guide, Part 1*. Washington, D.C.: U.S. Government Printing Office.
- Uyeki, E. S.
1964 "Residential distribution and stratification." *American Journal of Sociology* 69:491-8.
- Zelder, R. E.
1970 "Racial segregation in urban housing markets." *Journal of Regional Science* 10:93-105.

UNDERSTANDING THE STANDARDIZED INDEX OF DISSIMILARITY: REPLY TO MASSEY

Massey's comment indicates that there appears to remain a number of misconceptions about our two previous articles (Cortese et al., 1976; Cohen et al., 1976) regarding corrections in the calculation and use of the index of dissimilarity (D). We are currently in the process of preparing a more definitive treatment of the use of D in comparing cities and welcome this opportunity to address some of these misunderstandings.

The main point of our earlier work is that D is an inaccurate measure when one is comparing cities which differ in their proportions of the minority. This includes comparing the same city over time where the proportion of the minority has changed during that time or comparing occupational categories or any other unit of analysis when the proportions of the minority differ among categories (Farley, 1977; Simkus, 1978). This issue was raised by others before us (Duncan and Duncan, 1955; Taeuber and Taeuber, 1965:202-16). Nonetheless, researchers have continually ignored this limitation and the magnitude of the problem it presents never has been demonstrated fully. In recent discussions, arguments against the use of our suggested correction (Z_D) somehow turn into arguments against the necessity to correct D at all. These discussions thus miss the major point.

We maintain that the necessity to correct D for comparative studies exists and that our correction (Z_D) does allow for a comparative index of segregation. In fact, Massey has demonstrated very nicely with his data the degree to which the problem exists. His Table 1 lists nine counties. Looking only at the D scores, one would assume Somerset is the most segregated and Bergen the least segregated. Massey, in fact, has rank-ordered his nine counties according to D scores which at least implicitly acknowledges the propensity for comparative analysis across cities using D. Quite explicitly, he compares the rank-orderings of D scores and Z_D scores expecting that if our argument were correct, there would be a high correlation between the two. This mistake has been made previously by the Taeubers in their comment on our original article (Taeuber and Taeuber, 1976:885) and we dealt with it in some detail in our reply (Cohen et al., 1976:891). We understand the institutionalization of D and the sacredness it has acquired in sociology. However, expecting to validate a measure (Z_D) designed to correct the inaccuracies of another (D) by testing their correspondence is improper. The resulting lack of rank-order correlation between D and Z_D found by Massey is due to the fact that *D is an inaccurate index for such rankings.*

The Utility of Z_D

The proper interpretation of Massey's Table 1 is that, given the particular composition and density of each county, D is the percent of the minority, however large, that must be redistributed to achieve evenness. $D - E_D / 1 - E_D$ is the percent of the minority which needs to be redistributed in order to achieve a random dis-

tribution (Winship, 1977:1063-4). Z_D is the proper statistic for rank-ordering the counties according to their relative degree of segregation given their unique compositional characteristics (proportion of minority $[q]$, number of minority in a tract $[N_i]$, tract size $[T_i]$, and the degree of homogeneity of tract sizes).

Thus, a given Z_D is interpretable when used, as it should be, in comparison with another Z_D . A single Z_D tells little of value (i.e., the distance from randomness in units of the standard deviation) and, obviously since Z_D is a comparative index, would not be calculated when the researcher is interested in only one city. However, almost all uses of D involve comparisons: Is one school more segregated than another? Is one occupation more segregated than another? There is no need for confusion about "substantive vs. statistical significance." By suggesting the use of Z_D we are not interested that a researcher demonstrate that a large D is statistically significant but that one city can be judged to be more or less segregated than another.

Controlling for Relevant Factors

The admonition that Z_D "in no way removes or controls for [the] effects" (Massey, 1978:587) of q and T_i is simply wrong. Control includes methods other than holding constant. The fact that these two factors "vary from city to city" of course results in Z_D being expected to vary. That is the point. Z_D is criticized because it does not "remove or control for" changes in these two factors (Massey, 1978:587) or because it is "even more sensitive to differences" in composition (Simkus, 1978:83). The value of Z_D is that it is sensitive to these relevant factors but sensitive in a systematic and measurable way (unlike D), and that a proper measure for comparing cities must, as Z_D does, incorporate and account for differences in composition. q , T_i , N_i , and T are variables which must be measured and included in any calculation where they vary. Within one city q , T_i , N_i and T are constant but across cities they vary. Therefore, these are important variables which cannot be ignored. Our formula treats these factors as variables while D ignores their presence and is thereby distorted. Our point, again, is that Z_D is important when comparing cities and relatively unimportant when describing a single city. In such cases, D is sufficient for determining the proportion of the minority which must be redistributed without replacement to achieve evenness, even though evenness may be impossible empirically due to the compositional characteristics of that city.

Approximating the Hypergeometric Distribution

In our previous articles, we presented the correct sampling distribution of the index of dissimilarity—namely, the hypergeometric distribution. Like others (Winship, 1977; Taeuber and Taeuber, 1976), we assumed that this distribution would be approximated quickly by both the binomial and the normal distributions. This especially seemed likely since census tracts rather than blocks were used as the areal unit of analysis and the numbers (T_i) in a census tract tend to be large (average size 4,000). The idea that the expected D (E_D) would be small when T_i was large has led some to conclude that Z_D is unnecessary when T_i is large. However, this assumption is based on calculations using binomial approximations. We have discovered that this approximation is inadequate because of the problem of unequal tract sizes. To use the approximations, tract sizes must be both large and equal. Some tracts have a $T_i=0$, while others have many thousands. Therefore, while the hypergeometric is the exact distribution, the normally assumed approximations do not converge within significant digits. This means that the hypergeometric must be calculated for situations with large numbers. Like Massey, we discovered that this calculation was onerous even with large computer facilities. We had not anticipated the scale of this problem and are currently at work on a sensible resolution. However, while this computation remains difficult for the moment, we strongly urge researchers in the area to be sensitive to the issues we raise and not attempt to dismiss our substantive points on an issue of practicality which can and will be resolved.

Summary

We remain open to suggestions and alternative measures which would allow for the proper comparison of cities along the dimension of relative segregation. Massey believes he has shown logically that Z_D is inappropriate because it does not control for such factors as q and T_i . His notion of control seems to be making these factors constants. In reality, they are relevant variables and as such must be included in any calculation of relative segregation. Logically, therefore, Z_D is more appropriate than is D for intercity comparisons.

Massey further states that Z_D is inappropriate on the basis of empirical evidence. His evidence is the lack of a rank-order correlation between a nonrankable measure (D) and a rank-order index (Z_D). Besides the illogical nature of this attempted correlation, we repeat

(Cohen et al., 1976:891) that because one city among many has the largest proportion of minority to redistribute to achieve evenness, it does not necessarily mean that city is the most segregated. This is simply because the proportion to be redistributed is one thing and relative segregation is another. Relative segregation is dependent upon such factors as q and T_1 . The degree to which D is an inappropriate measure of relative segregation is evidenced in the low correlation between D and Z_D .

Massey also suggests the utility of Z_D may lie in being a "rough guide" in establishing the statistical significance of a D . He is correct in stating that such a task is unnecessary. We have never suggested the use of Z_D for that purpose but only for the purpose of providing a standardized index which is more appropriate for comparisons.

Some of our comments here have reiterated, perhaps unnecessarily, points we made in our reply (Cohen et al., 1976) to the Taeubers' (1976) comment. But neither of these articles is mentioned by Massey. We can only assume that either he has not found that debate helpful or that he has not found that debate.

Charles F. Cortese
R. Frank Falk
Jack Cohen
University of Denver

REFERENCES

- Cohen, Jack, R. Frank Falk and Charles F. Cortese
1976 "Reply to Taeuber and Taeuber." *American Sociological Review* 41:889-93.
- Cortese, Charles F., R. Frank Falk and Jack Cohen
1976 "Further considerations on the methodological analysis of segregation indices." *American Sociological Review* 41:630-7.
- Duncan, O. D. and Beverly Duncan
1955 "A methodological analysis of segregation indices." *American Sociological Review* 20:210-7.
- Farley, Reynolds
1977 "Residential segregation in urbanized areas of the United States in 1970: an analysis of social class and racial differences." *Demography* 14:497-518.
- Simkus, Albert A.
1978 "Residential segregation by occupation and race in ten urbanized areas, 1950-1970." *American Sociological Review* 43:81-93.
- Taeuber, Karl E. and Alma F. Taeuber
1965 *Negroes in Cities*. Chicago: Aldine.
- 1976 "A practitioner's perspective on the index of dissimilarity." *American Sociological Review* 41:884-9.
- Winship, C.
1977 "A reevaluation of indexes of residential segregation." *Social Forces* 55:1058-66.

THREATS TO INTERNAL VALIDITY IN BURGESS AND NIELSEN'S SOCIAL EXCHANGE EXPERIMENT

(COMMENT ON BURGESS AND NIELSEN, ASR JUNE, 1974)

In their article on interpersonal exchange, Burgess and Nielsen (1974:441) indicate that the deficiency of previous studies is that they

have been characterized by the exchange of work or time for money, with explicitly stated conditions. This practice, unfortunately, violates the definition of social exchange as a transaction of services or goods, the terms and timing of which are left unstated.

Recent experiences that we have had with the Burgess-Nielsen exchange paradigm lead us to believe that their research also suffers from this problem. We contend that certain seemingly innocuous features of their experimental apparatus provide the participants with a plethora of stimuli and cues which not only explicitly spell out the conditions of exchange, including the costs or gains from engaging in this activity, but in some cases appear to evoke such strong demand factors for exchange that one would be quite surprised if anything but exchange did occur.¹ In stating that "variation in our two independent variables, (1) resource value and (2) the availability of alternatives, determined whether exchange occurred," the authors, along with numerous references to the reinforcement contingencies or schedules associated with exchange, imply that performance in their study and the emergence of exchange was a consequence of subjects actually *experiencing* the experimentally arranged contingencies. For example, the contingency ER 3:3/IR 9:9² produced more exchange than the ER 9:9/IR 3:3 condition since the former provided greater rewards from exchange than from individual work. It is our belief, however, that the overriding influence of discriminative cues associated with the laboratory setting was so great that there was very little acquisition or exploratory experiencing of the schedules of exchange and individual work alternatives. In fact, the rapid acquisition and subsequent

¹ This comment also is methodologically opposed to Newport's (1975:840) criticisms of the Burgess and Nielsen study. Newport favors increasing the number of information cues—a move which we believe would confound the analysis even more. However, further studies should systematically vary information to assess its effect on exchange.

² The first ratio indicates the Exchange Ratio for subject A and subject B, respectively; the second indicates the Individual Ratios for these subjects (see Burgess and Nielsen, 1974:431).

maintenance of exchange or its reacquisition under advantageous conditions may be more a consequence of very explicit statements about the relative gain achieved through exchange, along with other reinforcing stimuli which are extraneous to the exchange itself. In this commentary we hope to explore in detail features of the Burgess-Nielsen experimental setting which inadvertently challenge the internal validity (Campbell, 1969:351-82) of their findings. To do this, we must first review their method for operationalizing exchange relations.

While other exchange paradigms have been employed (e.g., Michaels and Wiggins, 1976), Burgess and Nielsen operationalize exchange according to a traditional operant model. Two partners operate separate consoles, each with a simple exchange button response (i.e., pressing a PAY OTHER button) or a work alone alternative (i.e., pressing a PAY SELF button). These concurrent operants (Ferster and Skinner, 1957: 703-21) provide elegant and straightforward measures of exchange and nonexchange over time. In spite of the attractive features of this design, we soon discovered that certain aspects of the apparatus created problems. One of these is the display counter indicating the work requirement for successful exchange. This counter shows the exchange ratio, or the number of responses a subject must make on his PAY OTHER button to give his partner .01¢. In addition, the subject's own individual ratio is given below the PAY SELF button, though not his partner's. This labeling of individual and exchange ratios proves to be very salient for subjects; it informs them that they can, in most conditions, make more money, at a faster rate, by exchanging instead of working alone. Though Burgess and Nielsen argue that the relative values of exchange and work alone payoffs largely determine exchange in their study, their own data suggest that the ratio information accounts for the observed results. While Burgess and Nielsen (p. 433) emphasize that rates of exchange in Experiment I were low during the initial baseline condition, the data in their Figure 2 actually show considerable intergroup variability. Some groups (e.g., groups two, three, ten and eleven) obtained steady state exchange at the 90% level following a brief exposure to the ER 3:3/IR 9:9 contingency. Group six, which demonstrated less than 10% exchange in the first sixty minutes of baseline, jumped to 100% exchange at the beginning of a second baseline session. These data hardly support the Burgess-Nielsen contention of "step-by-step" acquisition. One must ask why it took anyone very long to begin

exchanging when it was so obviously profitable to do so. Unfortunately, there is no information concerning the debriefing of experimental subjects, so we can only guess what they were thinking. It is not inconceivable that what little exploratory behavior did occur may have been motivated by disbelief that this was all there was to the experiment, that it was really that simple, and that the payoff ratios given on the display counters were actually in effect. What is also remarkable here is that with a shift to a less favorable condition for exchange (ER 9:9/IR 3:3), there is an instantaneous readjustment, with steady state nonexchange obtained by almost all groups within the first fifteen minutes of this new condition. Finally, high probabilities of exchange are quickly reacquired for seven of the ten remaining groups following a return to the original baseline condition. To us, this behavior looks suspiciously like rule governed performance instead of behavior arising purely from the experienced schedules of reward, i.e., contingency-shaped performance. Providing detailed information on the payoff schedules not only increases subject's awareness of the contingencies but is a salient determinant for exchange. Technically, the description of contingencies functions as a discriminative stimulus (Skinner, 1969: 138), informing subjects when it is advantageous to exchange and when it is not.

Other evidence supporting the notion that exchange in this study may in large part be attributed to contingency descriptions can be seen in Experiment II. For two of the six groups receiving conditions C (ER 3:3/IR 9:3) and D (ER 3:3/IR 9:2), there is an immediate reversal in the probability of exchange with transition to condition D. Considering the slight value change in the individual work alternative for subject B, there should be a more gradual decline in exchange if subjects were merely under the control of reinforcement contingencies. At least one subject in group three at times does actually show this gradual change. In summarizing this second experiment, the results suggest that different subjects were governed by diverse aspects of the experimental arrangements. Some behaved in terms of scheduled contingencies, while others were affected mainly by descriptions of the payoff ratios. The main point is that Burgess and Nielsen cannot evaluate the impact of a change in reinforcement contingencies per se, since subject performance was altered in part by the descriptive information. Skinner (1969) convincingly argues that control by contingencies is theoretically and empirically different from control by descriptive rules. It is abun-

dantly clear that the payoff ratio information is an independent source of variation confounding the present results and their interpretation.

We find it peculiar that, on a theoretical level, Burgess and Nielsen argue that social exchange, in contrast to economic exchange, typically is characterized by a distinct lack of information and rules (see p. 441). Presumably, this means that instructions, contracts, rules or other verbal cues are not legitimate variables in the analysis of social exchange. It is therefore surprising that these researchers have built so many informational stimuli into their experimental setting. While they may be correct in pointing out that social exchange may not be guided by explicit rules, implicit norms (such as the desirability of reciprocating gifts received from someone else or the maxim "it is better to give than receive") may function in a variety of exchange encounters. Not only might these norms have tremendous practical significance in maintaining social exchange, but also may prove significant in formulating a viable theory of social exchange. Therefore, contrary to Burgess and Nielsen, we believe that rule governed social exchange is a legitimate avenue for research. In fact, the results of the present experiments suggest that rules and descriptions may be powerful variables. The only problem is that the descriptions and actual contingencies are confounded in the Burgess and Nielsen research.

Additional features of the experimental setting also raise both methodological and theoretical difficulties. A SELF TO OTHER counter allowed subjects to monitor how much they had given their partner. A second counter, OTHER TO SELF, indicated what a subject had received from his partner through exchange. Also, a blue signal light below the OTHER TO SELF counter was actuated every time one's partner pressed his PAY OTHER button. Since a given ratio of PAY OTHER presses were required to generate a point for the other person, the blue light bridged the time delay between the pressing and the onset of the point (p. 433, fn. 8).

Oddly enough, probably the most damaging feature of the apparatus is this blue light. In providing the blue response light, we have in effect eliminated the essence of social exchange, which is a temporal delay in reciprocating what another has given us. The interesting thing about social exchange is that, in real life, we do not necessarily see our exchange partners at work paying off their debts to us. It may be critical to know that the pauses and delays associated with the generating of exchange rewards may at times jeopardize the exchange or make it difficult to emerge in the

first place. Unfortunately, the blue light in the Burgess and Nielsen study deprives us of this knowledge. We began to appreciate the problems it creates during our own research when we focused on the effects of fixed and variable reward schedules on the acquisition of exchange.

It is well established that fixed-ratio schedules of reinforcement generate distinctive individual response patterns when compared to performance on variable-ratio schedules (Fester and Skinner, 1957). Thus, we expected fixed and regular patterns of reward during exchange to affect subject's behavior differently than when rewards occurred on a variable basis. To study this, we constructed an apparatus identical to Burgess and Nielsen's, though we omitted any description of exchange and individual work ratios. A pretest found no change in the pattern of exchange as a function of fixed or variable reward schedules. Also there were no disruptions in responding in alternating between fixed and variable exchange contingencies. During debriefing, subjects reported that the blue signal light, indicating the PAY OTHER pressing of their partner, was more important than the points on the OTHER TO SELF counter.

It became apparent that the blue light was a critical stimulus in strengthening exchange performance. While the signal light had been added as a temporal bridge, it clearly had other functional properties. As a cue, the light signaled a willingness to cooperate and, hence, influenced the selection of exchange over the individual alternative. More importantly, it served as the most immediate consequence to exchange responses. The blue light, which subjects described as "effort" or "trying" by their partner, is a strong social reinforcer supporting exchange activity. The effects of the reward schedules involving the counter points often were not observed because of the reinforcing power of the signal light. This explains the lack of an observed difference in the exchange pattern for fixed vs. variable schedules of points. Rather, the signs of effort, which were almost continuous when the partner elected to exchange, determined the rate and pattern of exchange responses. This analysis is given added confirmation with the removal of the light signal for subsequent experimental dyads. Under these conditions discernible differences in fixed and variable reward schedules were observed. The salience of these observations for the Burgess and Nielsen results are obvious. It is held that exchange responses of subjects in their research also were strengthened by the signal light, making it difficult to assign the observed effects only to resource value and

availability of alternatives. Also, as we have stated, the blue light may be questioned on theoretical grounds since it eliminates the pauses and delays often associated with social exchange.

The SELF TO OTHER and OTHER TO SELF counters are, however, theoretically important conditions for social exchange, and the comparison between these counters enables subjects to monitor the degree of equity in the exchange. Since Burgess and Nielsen are very much concerned with equity, the presence of both counters seems reasonable. Our major point here concerns the separate and combined effects of these counters and their influence in the experimental setting. While Burgess and Nielsen fail to provide a rationale for these counters, their article implies the monitoring function of these devices. Thus observations of equal monetary payoffs are attributed to equity concerns. However, subjects in our research have reported a tendency to compete with their partner. High rates of giving by both partners, resulting in equal payoffs, may be due to each subject attempting to keep a positive difference between the SELF TO OTHER and OTHER TO SELF counters. Equitable payoffs may not reflect, therefore, equitable behavior but rather a process of "one-upmanship" maintained by (a) the escape from the aversive properties of "being in debt," indicated by a deficit between SELF AND OTHER counters; (b) the positive reinforcement value of "having the other in debt," indicated by a positive difference between SELF and OTHER counters; or both of these processes.³ Burgess and Nielsen instituted a timer on the PAY OTHER button, equalizing the rate of response of the partners. This device perhaps diminished the competition process we have outlined, but the emergence of competition in exchange is an important process which must be distinguished from equity processes. Exactly how the counters functioned in the Burgess-Nielsen experiments can not be ascertained from the results as presented.

In addition to the combined effects of the point counters, it is necessary to recognize their separate effects for social exchange. It is obvious that the monetary payoffs registered on the OTHER TO SELF counter constitute a unique independent variable. The level of exchange activity on the PAY OTHER button would seem related directly to the points re-

ceived. Actually, it is these points which operationally define resource value in this study. However, a study by Weiner (1977) suggests that, at least for some individuals, giving to other may be influenced by the points on the SELF TO OTHER counter. Weiner shows that signs of giving (e.g., points indicating other has received a monetary payoff) themselves can function as reinforcers which strengthen and maintain giving (e.g., PAY OTHER button pressing). Thus the SELF TO OTHER points constitute a separate variable which itself may account for exchange activity. The separate and combined effects of the point counters could operate in the Burgess and Nielsen experiments and it is not obvious which effects contributed to the observed behavior of the subjects. Subsequent research must recognize these distinct effects and control for the effects of signs of giving when investigating resource value and exchange transactions.

In this section we have outlined the possible confounding variables operating in the Burgess-Nielsen experiments. We recognize the difficulty of controlling the influence of extraneous variables in experimental work; but when sources of variation are identified by replication and extension, this must be communicated to interested researchers. Burgess and Nielsen provide an extremely useful setting for analyzing social exchange and, by specifying the operating variables in their setting, we hope to refine the analysis and investigation of social exchange. In this regard, the subsequent part of this comment notes briefly some design problems and difficulty in presentation of the dependent measures of the Burgess and Nielsen study which constitute further considerations for research on exchange relations.

Design and Measurement Problems

Burgess and Nielsen employ an intrasubject replication design. Such a design seems appropriate for research where the interest is focused on a variable's operation over time. Glass et al. (1975:33) uphold the viability of these kinds of designs in social research but also note some difficulties. Of specific interest is the problem of the "multiple-intervention inference" (p. 58). Problems of invalidation occur when two or more interventions, i.e., treatments, are applied to the same time-series. Glass et al. (1975:58) state that "in such cases, the second or later interventions can produce effects which are unique to experimental units exposed to preceding interventions." In the Burgess and Nielsen experiments all exchange dyads initially were exposed to balanced contingencies (e.g., ER 3:3/IR 9:9) and then to

³ Subjects in our study often indicated that they wanted to give the other person more than they received and there was a perceived loss of esteem if one could not keep up with another's exchange.

imbalanced arrangements (e.g., ER 3:3/IR 9:3). It is possible that the disruption of exchange in the imbalance condition is due to the previous exposure of subjects to balanced contingencies and is not a function of the unequal structural opportunities. The order of intervention therefore strongly suggests an alternative explanation for the Burgess and Nielsen findings. This is even more a possibility when one considers that the same dyad experienced five or more separate interventions. It is recommended that researchers attempt to control for the effects of order of treatment in further exchange studies employing operant or other time-series designs. A counterbalanced design can control for the multiple-intervention problem. When the experiment consists of two interventions I_1 and I_2 , one experimental unit should be exposed to I_1 and then I_2 , while a second unit received the reverse order of I_2 and then I_1 . Such counterbalancing becomes practically unfeasible as the number of interventions applied to a single unit increases (see Glass et al., 1975:35). This suggests that a researcher should opt for an unambiguous assessment of a few critical variables, i.e., one or two treatments, than for a confusing assessment of multiple interventions.

A final concern is addressed to some aspects of the data presentation and measurement in the Burgess and Nielsen study. A major portion of the data are reported in graphic form for each experimental unit. However, in following the same dyad over the various experiments it is apparent that data either are omitted or rearranged in nontemporal order. For example, group seven is initially treated with condition A, then condition B and back to the A condition during experimental sessions one through three. The next data reported for this group appear in Experiment 5 with the application of condition G during session fourteen. But there is omission of the experimental treatments applied between sessions three and fourteen. The reader is given no explanation of such omissions and this works against the credibility of the entire experiment. In addition, the data presented for group ten, comparing condition A with B, are actually not a temporal sequence of exposure to these treatments. Graphic portrayal of Experiment 1 for this group compares session one (condition A) with session seventeen (condition B). The observed change for the B condition is not presented accurately since condition A did not actually precede the B treatment. Such data analysis is confusing and misleading to the reader.

Finally, it is necessary to comment on the use of relative frequency of exchange, or probability of exchange, as opposed to frequency

or rate measures. The Burgess and Nielsen measure reflects the extent to which subjects choose exchange over the work alone alternative. For example, in the present study a subject who displays ten responses in exchange and zero responses in the alternative has a relative frequency of 1.00. But so does the subject making 120 exchange responses and zero responses in the alternative. If we are interested in the effects of contingencies on exchange it is clear that the relative measure hides essential information. Obviously investigators are not interested merely in the individual's probability of selection of exchange over alternative activity. It makes a substantive difference if an individual emits ten rather than 100 exchange responses during a given time period. Rate of exchange reflects the "work" that an individual puts into the exchange relation and as such it is an essential measure in analysis of exchange. Burgess himself has extolled the virtues of rate of response as a dependent variable (Burgess and Bushell, 1969:166). Proportion data can be deceptive when depicting change over time if the base is changing in value. Thus a subject who emits 100 exchange responses and 100 alternative responses has a relative frequency of 0.50. If his exchange responses stay constant at 100 but his alternative responses decline to 50, then the relative frequency is 0.67. The reader might conclude that there has been some improvement in exchange activity when, in fact, it has not changed at all. Given these considerations and the substantive concerns of exchange theory, it is advisable to present rate of response data which may be supplemented by relative frequencies.

Conclusion

This comment has demonstrated that variables other than resource value and available alternatives were operating in the Burgess and Nielsen experiments. The effects of contingency descriptions (individual and exchange ratio) and signs of effort (the blue light) constitute alternative hypotheses for their observations. Also, we show that the signs of giving (SELF TO OTHER counter) and points received (OTHER TO SELF counter) constitute separate independent variables for a theory of social exchange. The combined effect of these conditions may generate equitable behavior or competition, processes which are not distinguished in the present research. Finally, our comment identifies the problem of multiple-intervention associated with the Burgess and Nielsen design and suggests how to improve

data presentation in subsequent experimental analyses of exchange relations.

W. David Pierce
W. Andrew Harrell
University of Alberta

REFERENCES

- Burgess, R. L. and D. Bushell, Jr.
1969 "The experimental analysis of social process in the field." Pp. 143-74 in R. L. Burgess and D. Bushell, Jr. (eds.), *Behavioral Sociology: The Experimental Analysis of Social Process*. New York: Columbia University Press.
- Burgess, R. L. and J. M. Nielsen
1974 "An experimental analysis of some structural determinants of equitable and inequitable exchange relations." *American Sociological Review* 39:427-43.
- Campbell, Donald T.
1969 "Artifact and control." Pp. 351-82 in R. Rosenthal and R. Rosnow (eds.), *Artifact in Behavioral Research*. New York: Academic.
- Ferster, C. B. and B. F. Skinner
1957 *Schedules of Reinforcement*. New York: Appleton-Century-Crofts.
- Glass, G. V., V. L. Wilson and J. M. Gottman
1975 *Design and Analysis of Time-Series Experiments*. Boulder: Colorado University Press.
- Michaels, J. W. and James A. Wiggins
1976 "Effects of mutual dependency and dependency asymmetry on social exchange." *Sociometry* 39:368-76.
- Newport, F.
1975 "An experimental analysis of some structural determinants of equitable and inequitable exchange relations (comment on Burgess and Nielsen, ASR June, 1974)." *American Sociological Review* 40: 838-43.
- Skinner, B. F.
1969 *Contingencies of Reinforcement: A Theoretical Analysis*. New York: Appleton-Century-Crofts.
- Weiner, H.
1977 "An operant analysis of human altruistic responding." *Journal of the Experimental Analysis of Behavior* 27:515-32.

REPLY TO PIERCE AND HARRELL

In a reply to an earlier comment on the Burgess-Nielsen social exchange experiments, we remarked that there were several features of our study that might lead to misunderstanding and confusion. One of those features was that the study is one of the first to analyze direct exchange relations experimentally. As we noted in our article, the sociological study of exchange processes has been marked by many more words than data. Thus, ever since

the publication of our article in the *Review* I have been eagerly awaiting the publication of other experimental research on exchange behavior. I have waited in vain. Indeed, our research seems to have generated only more words.

Admittedly, the empirical analysis of direct exchange relations is difficult. This is the case primarily because in direct exchange relations the actors themselves control the actual distribution of rewards. As we pointed out (Burgess and Nielsen, 1975), this leads to considerable complexity in experimental design as well as in data analysis. Regrettably, Pierce and Harrell seem to have been overwhelmed by the complexity. At the risk of reinforcing further comments on our research in place of careful experimental research, I will take this opportunity to correct as best I can the misconceptions that might arise from their Comment.

Economic Exchange

Pierce and Harrell claim that our experiments are more relevant to rule governed economic exchange than to the social exchange that is of more interest to sociologists. The basis for their assertion is that the experimental apparatus we used explicitly informed the subjects in our experiments what the prevailing contingencies were and that the subjects were responding to the normative features of those schedules rather than to the reinforcers they actually received. They indicated that it is their "belief . . . that the overriding influence of discriminative cues associated with the laboratory setting was so great that there was very little acquisition or exploratory experiencing of the schedules of exchange and individual work alternatives" (pp. 592, emphasis added).

Their belief, however, is based upon some pretty shaky reasoning. For example, they are disturbed by what they feel is unusually "rapid acquisition and subsequent maintenance of exchange and its reacquisition under advantageous conditions . . ." (p. 592-3). Yet, in the same paragraph they appear, in one breath, to be perplexed that our subjects responded so quickly to the prevailing contingencies and, in the next, by the fact that, in some cases, it took so long for them to come under the control of the contingencies. It is difficult, indeed, to respond to arguments such as these. Nonetheless, they seem to decide that they are most concerned by the short acquisition periods and this leads them to conclude that the "behavior looks suspiciously like rule governed performance instead of behavior arising purely from the experimental schedules of reward, i.e., contingency-shaped performance" (p. 593).

Had they read our article carefully, they would have noted many instances where the subjects' differential performances clearly indicated that they were responding to the actual reinforcers received from their partners. For example, in our third experiment, we noted that the more powerful members of the exchange relation (those with more attractive alternatives) refused to exchange unless their partners increased their rewards to them (p. 437). Analysis of our polygraph data, which unfortunately were not included in our published account, indicate further that the subjects were responding in a contingent fashion to how much their partners were rewarding them.

Moreover, in another experiment (Nielsen, 1974) we examined the effect of seeing or not seeing the Exchange Ratio counter that displays the prevailing schedules of reinforcement for exchange. The results from that experiment indicated that it essentially made no difference. Admittedly, it could make a difference in the acquisition of the exchange behavior but we were not particularly interested in acquisition *per se*.

Pierce and Harrell also confuse Skinner's analysis of rule governed behavior with the use of discriminative stimuli associated with reinforcement schedules. Most operant experiments include the use of various discriminative stimuli (e.g., red and green lights) associated with various experimental conditions (e.g., extinction vs. reinforcement). This is especially the case where the investigator is not trying to determine whether the behavior under investigation is an operant or not, but instead is exploring the effects of diverse schedules. This is what we were doing and since our subjects were verbal, we used real numbers rather than red or green lights.

In sum, it is the free operant characteristic of our research setting that led us to say that the setting fit the definition of social rather than economic exchange. While the subjects, indeed, were told what the exchange schedules of reinforcement were, the critical controlling variable was how much their partners actually rewarded them. Again, inspection of our polygraph records reveals that subjects would often reduce their rate of exchange behavior until their partners increased their rate. Thus, an attractive schedule of reinforcement was not enough. Nor was it sufficient if one's partner was only exchanging half heartedly. Clearly, the probability of exchange was a direct function of the amount of reinforcement received compared to the alternatives. If Pierce and Harrell have any real doubt that our subjects were responding to the reinforcements

received, I suggest that they do the following. First, read our article carefully. If that does not do the job, then they should read our reply to Newport. If they remain unconvinced, I suggest that they attempt a direct replication of our experiment.

Social Exchange

After arguing that our subjects may have been responding to the demand characteristics of the schedule information presented on the exchange consoles, they then argue that the subjects may have been responding instead to the blue light we used to indicate to each subject that their partner was working for them (see Burgess and Nielsen, 1974:433). We included this design feature to approximate some of the features of face-to-face interaction where each actor is simultaneously sending and receiving information.

Pierce and Harrell's reasoning here is as dubious as it was in the preceding section. For example, they refer to some pretesting which they did where their subjects purportedly said that the blue light going on and off was more important to them than the points they received from their partner. I do not know what to make of this for they present no data nor do they say whether the points received were exchanged for money; and, if so, whether the amounts of monetary pay were equivalent to ours. As we noted in our original article, our subjects were hired from the university's employment office. These volunteer subjects were fully informed that the only pay they received was to be based on their actual earnings in the experiment. Thus, our subjects were highly motivated and the reinforcer was money. Indeed, in questionnaires distributed at the conclusion of our experiments, our subjects unanimously noted that their behavior was motivated chiefly by the money they received. Apart from their verbal responses, a close examination of our data will indicate that to be the case.

Pierce and Harrell also base their argument on the fact that in their pretest their subjects did not respond differentially to fixed vs. variable ratio schedules. They then claim that when they removed the blue light, their subjects displayed "discernible differences" (p. 594) when these two schedules were compared. Three things bother me here. First, what are discernible differences? Again, they present no data. Second, if, as they claim, they "constructed an apparatus identical to Burgess and Nielsen's" (p. 594) there would be no way that they could differentiate fixed from variable ratio reinforcement. Indeed, as noted in our article (p.

432), our apparatus employed a mixed ratio-interval schedule. We did this to prevent the emission of responses greater than two per second. Thus, we were able to control for differential dexterity on the part of our subjects and to place an upper limit on the amount of reinforcement that could be received. We were, after all, using real money. In any event, given this design feature, I would like to know how Pierce and Harrell were able to differentiate fixed from variable ratio schedules. My suspicion is that their apparatus was not, in fact, like ours at all. Moreover, I doubt that they were using monetary schedules equivalent to ours.

Third, I see no reason why fixed vs. variable schedules should have any effect at all in direct exchange. After all, each subject's reinforcement is governed by the *other* person's behavior. That is the way exchange is defined and it is the way our subjects behaved. Now, if one actor's behavior actually was being maintained by his or her delivery of money to the other actor, then we would expect to see different response patterns for fixed and variable reinforcement schedules. But, of course, then we would not be studying exchange at all but something akin to altruism. I suggest they have much more pretesting to do.

Design and Measurement Problems

Pierce and Harrell raise a number of questions concerning our use of a reversal design. Most of their arguments here seem to be either beside the point or further reflections of their failure to read our article carefully and to consider our experiments in their entirety. For example, their concern for "multiple intervention inference" (p. 595) is misplaced. Actually, we obtained the same kinds of responses to our experimental conditions regardless of what the preceding experimental conditions were. Indeed, the conditions we explored were so powerful that they overruled the effects of previous experimental histories. In this connection, they seem troubled by the data presented for group ten in Experiment 1. The fact that this group displayed behavior in condition B similar to the other groups despite the intervening presentation of other conditions adds to the strength of the results rather than detracts from them as they suggest.

The other experiments that were not reported in our article dealt with questions that were not related specifically to equity-inequity.

They dealt with such features as the effect of seeing or not seeing the Exchange Ratio counter and the effects of having the opportunity to take money from each other. All our data could not be presented in one report.

Finally, Pierce and Harrell criticize us for using relative rates of response rather than absolute rates. They should read the concurrent operant literature—we were simply following the conventions. Had we presented absolute rates, the differential effects of the various experimental conditions would have remained the same; they simply would have been more difficult to communicate. In any event, our primary interest was in the relative frequency of exchange given the availability of an alternative response.

Conclusion

I have tried to show that the criticisms of our study by Pierce and Harrell are unwarranted. Most of their criticisms were based on the failure to read our article carefully or to their failure to understand it. I must admit that I am perplexed by their writing their comment in the first place. It would seem to me that their first order of business would be to attempt to directly replicate our experiments. Upon doing that, they then may be able to demonstrate experimentally, rather than verbally, that certain conditions need to be specified more clearly than we did in our initial experiments. Were they to do so, I would be among the first to applaud. Given the complexity of exchange relations, we are in desperate need of more data. I should hope that Pierce and Harrell could contribute constructively to the solution of this problem.

Robert L. Burgess
Pennsylvania State University

REFERENCES

- Burgess, R. L. and J. M. Nielsen
1974 "An experimental analysis of some structural determinants of equitable and inequitable exchange relations." *American Sociological Review* 39:427-43.
- Burgess, R. L. and J. M. Nielsen
1975 "Reply to Newport." *American Sociological Review* 40:843-5.
- Nielsen, J. O. M.
1972 "Experimental analysis of equitable and inequitable social exchange." Ph.D. dissertation, University of Washington.

THE EFFECTS OF INTERNATIONAL ECONOMIC DEPENDENCE ON DEVELOPMENT: A CRITIQUE

(COMMENT ON CHASE-DUNN, ASR
DECEMBER, 1975)

In "The effects of international economic dependence on development and inequality: a cross-national study," Christopher Chase-Dunn (1975) utilizes panel regression analysis to support the hypothesis that the penetration of developing areas by foreign capital produces a "development of underdevelopment."¹ Chase-Dunn's universe consists of the countries with a GNP per capita of less than \$406 in 1955 (see Appendix below). The segment of his analysis of interest here is the regression of GNP per capita 1970 on GNP per capita 1950, percent of GDP in mining 1950, and a measure of per capita foreign investment dependence 1950-55. Chase-Dunn's findings (see Table 1, equation 1) show that, with GNP per capita 1950 and specialization in mining 1950 controlled for, there is a significant negative relationship between investment dependence and later development (GNP/capita 1970). The results appear to support the assertion of dependency theory that, in the aggregate, economic dependence hinders development. However, when reviewed in more depth, the study seems to raise additional questions rather than provide solutions. The main thrust of the following is to comment on Chase-Dunn's measure of economic dependence and to suggest a more valid alternative.

Measurement of Economic Dependence

In order to study the effects of investment dependence on development, Chase-Dunn uses a national accounting item, debits on investment income, as an index of economic dependence. This statistic "reports all profits made by foreign direct investment in the 'host' country (regardless of whether or not they are repatriated)" (Chase-Dunn, 1975:728). The contention is that this measure, reported in the International Monetary Fund (1959-63) balance of payments data, represents a good proxy for the stock of foreign capital in a country at a point in time. Using such a proxy, Chase-Dunn assumes we may infer the effects of economic dependence on development by comparing the size of the stock of capital held by foreign interests with changes in social and economic indicators for a sample of developing areas.

Chase-Dunn uses this measure of the stock of foreign capital because accurate data on foreign investment capital in developing countries are unavailable.² However, the precise interpretation of this proxy is that debits on investment income presents either a direct outflow of resources in the form of repatriated earnings and interest payments, a potential for such an outflow, or, in most cases, a combination of both. It should be generally agreed, by even the most ardent critics of dependency theory, that to the extent that such resource drains occur, they should retard growth. In fact, the governments of most developing countries, cognizant of this problem, impose strict controls on flows of foreign exchange. Viewed in this context, the Chase-Dunn paper tells us nothing new. Rather, the study supports the accepted notion that resource outflows will impede social and economic development.

A second problem with the choice of debits on investment income as a measure of economic dependence is that the variable itself may be distorted severely. Depending on various tax considerations, both at home and in the host country, coupled with varying corporate objectives, there exist vast incentives for firms to overstate or understate reported profits. The tendency is facilitated through the use of accepted accounting techniques including accelerated depreciation, inventory valuation manipulations, and in the case of extractive industries, depletion allowances. Because such incentives (and the means to carry them out) exist, official profits data may be distorted in one direction or another depending on the country in question. The results of cross-sectional regressions using such a variable would be difficult to interpret.

Few theorists would deny that the degree to which a nation is dependent on foreign capital is a function of the extent to which such investment is profitable. However, a third criticism of Chase-Dunn's dependency measure is that many investment projects have long gestation periods with large capital requirements. In addition, extended periods of time frequently are required before such projects become operational, and possibly many more years are required before they become profitable, if at all. Investment in what has been termed social overhead capital such as roads and power networks are classic examples of investment projects exhibiting such behavior. The debits on investment income variable fails to capture this

¹ This statement is initially attributed to Frank (1969).

² For the most concise data on stocks of foreign capital in developing nations see Organization for Economic Cooperation and Development, 1972.

crucial aspect of financial inflows, that profits may not reflect accurately the existing stock of foreign capital. It is conceivable that, if all foreign investment in a particular country went into such social overhead uses, the debits on investment income might indicate that no dependence on foreign capital exists.

An additional criticism of the Chase-Dunn analysis is that the specification of his model essentially relates a change in the flow of output (i.e., GNP/capita in 1950 to GNP/capita in 1970) to a stock of foreign capital. The theoretical justification for relating changes in a flow variable to a fixed stock variable is uncertain. Since the process of accumulating a stock of foreign capital presumably already has had an effect over preceding years, the level of GNP per capita in a given year should be correlated with the stock of foreign capital in that year, but changes in GNP per capita over subsequent years should be uncorrelated with past levels of the stock of foreign capital stocks. Economic growth literature avoids this problem by relating changes in output to *changes* in the capital stock (i.e., investment).

The question remains, What aggregate measure can be used to study the effects of the presence of foreign capital inflows on development? Chase-Dunn points out that an alternative balance of payments statistic, direct foreign investment, theoretically could provide a good measure of investment dependence, but that it exhibits too much yearly variance to be of any practical use. In addition, the procedures for measuring direct foreign investment vary from country to country (it is frequently estimated), and also tend to distort this variable.

A more useful index of international invest-

ment dependence is the balance on nonmonetary sectors capital. This balance of payments item summarizes five separate international accounts: direct foreign investment, private long and short term capital flows, and finally, local and central government capital flows. Fundamentally, this new variable measures the degree to which international capital flows into the host country. The statistic includes long and short term loans for private business, direct investment by multinational enterprises, and, to a large extent, the dependence of a central government on foreign capital.

Results

Table 1 illustrates the results of a study—similar to that performed by Chase-Dunn but using the balance on nonmonetary sectors capital—as an alternative measure of investment dependence. Specifically, the new formulation incorporates panel regression analysis where:

Y_t = Log GNP per capita 1971 measured in 1964 U.S. dollars (International Bank for Reconstruction and Development, 1973);

Y_{t-1} = Log GNP per capita 1960 measured in 1964 U.S. dollars (International Bank for Reconstruction and Development, 1971);

M = Percent of GDP in mining 1960 (International Bank for Reconstruction and Development, 1971); and

I = 1959–61 average investment dependence: the balance on nonmonetary sectors capital per capita measured in 1964 U.S. dollars (International Monetary Fund, 1959–63).

Table 1. Panel Regression Estimates of the Effects of Investment Dependence and Specialization in Mining on Economic Development

Equation 1 (Chase-Dunn, 1975:731, Table 2)				Equation 2 (Stumpp, Marsh, Lake)			
	b	S.E.b.	Beta		b	S.E.b.	Beta
Y_{t-1}	1.50	.16	1.05**	Y_{t-1}	1.061	.084	.88**
M (1950)	.12	.046	.28**	M (1960)	.007	.006	.10
I (1950–55 average)	-.187	.059	-.41**	I (1959–61 average)	.003	.005	.04
Constant = -.39				Constant = .165			
$N=25$						$R=.937$	
						$R^2=.878$	
						$S.E.=.248$	
$Y_t = \ln$ GNP per capita 1970				$Y_t = \ln$ GNP per capita 1971			
$Y_{t-1} = \ln$ GNP per capita 1950				$Y_{t-1} = \ln$ GNP per capita 1960			
M = Percent of GDP in mining 1950				M = Percent of GDP in mining 1960			
I = Investment dependence, debits on investment income, 1950–55 average				I = Investment dependence, balance on nonmonetary sectors capital, 1959–61 average			
* $p < .05$ ** $p < .01$							

Some care has been taken to obtain data from the same sources used by Chase-Dunn. Unfortunately, because capital account data for many of the 35 sample countries are incomplete or unavailable for 1950, the time span covered by this study, when compared to Chase-Dunn's, has been reduced from 20 to 11 years (1960 to 1971).

At the zero-order level, Ln GNP per capita 1971 has a Pearson product-moment correlation of .93 with Ln GNP per capita 1960, .46 with investment dependence 1959-61, and .38 with mining 1960. Note that these lagged relationships are all positive. The regression coefficients estimated by the new equation prove to be quite interesting. The most surprising result is that in Table 1, equation 2, contrary to Chase-Dunn's analysis, investment dependence is shown to have *no significant effect* on development. For the sample of 35 developing countries, when log 1960 GNP per capita is held constant, the presence of foreign capital inflows neither helps nor hinders the development process.

Because the inclusion of a lagged dependent variable may bias³ the coefficients in both equations of Table 1, Table 2 illustrates an alternative specification of the regression. Instead of regressing 1971 GNP per capita on 1960 GNP per capita, Table 2 (equation 3) regresses percent *change* in GNP per capita between 1960 and 1971 on investment dependence 1959-61 and mining 1960. At the zero-order level, both these prior variables have essentially equal, significant, positive relationships to the rate of change in GNP per capita. The latter variable has a .37 correlation with mining 1960 and a .32 correlation with investment dependence 1959-61. The multiple regression results for equation 3 are basically the same as for equation 2. Neither prior investment dependence nor prior specialization in mining has any significant effect on the rate of growth in GNP per capita between 1960 and 1971.

In addition to GNP per capita, Chase-Dunn (1975) used two other measures of the dependent variable, economic development. These were the percent of the labor force in nonagricultural occupations in 1960 and kilowatt hours of electricity consumed per capita in 1965. Chase-Dunn found that his measure of investment dependence had significant negative effects on these measures of development as well. To replicate as closely as possible this

Table 2. Alternative Specification of the Effects of Investment Dependence on Economic Development

	Equation 3 (Stumpp, Marsh, Lake)		
	b	S.E.b.	Beta
M (1960)	.013	.009	.278
I			
Investment dependence 1959-61 average	.007	.007	.178
Constant	-0.641		R=.400
N=35			R ² =.160
			S.E.=.392

Y_t=Percent change in GNP per capita between 1960 and 1971

M=Percent of GDP in mining 1960

I=Investment dependence, balance on non-monetary sectors capital, 1959-61

part of Chase-Dunn's analysis, we present Table 3 which estimates the effect of our substitute measure of investment dependence on each of these aspects of economic development. Once again, to avoid the bias of a lagged dependent variable, the dependent variable we use in equation 4 represents percent *change* between 1960 and 1970 in the percent of the labor force in agriculture, and in equation 5, the percent *change* between 1960 and 1970 in KWH electricity consumption per capita. Data were available (International Bank for Reconstruction and Development, 1976) for 32 of the 35 countries in Table 2.

Consider first the results for equation 4 in Table 3. With percent of GDP in mining in 1960 controlled for, investment dependence, 1959-61 has a beta of -.374 in relation to percent change in percent of the labor force in agriculture. Since we are measuring the percent in agricultural (rather than nonagricultural) occupations, a negative sign indicates economic development (i.e., a decline in the percent in agriculture). Thus, we find that among these less-developed countries, the more dependent a country was in 1959-61, the greater was its decline in the percent in agricultural occupations from 1960 to 1970. In other words, economic dependence had a significant *positive effect* on this labor force aspect of development. The result for equation 5, again with percent of GDP in mining controlled for, is that investment dependence has a negative, but nonsignificant, effect on change in electricity consumption per capita between 1960 and 1970. The important point, again, is that the

³ Specifically, the inclusion of a lagged dependent variable in the right hand side of a regression may cause its measured coefficient to be biased downward and the coefficients of other independent variables to be biased upward.

Table 3. Effects of Investment Dependence on Other Measures of Economic Development

Equation 4 (Stumpp, Marsh, Lake)				Equation 5 (Stumpp, Marsh, Lake)			
	b	S.E.b.	Beta		b	S.E.b.	Beta
M (1960)	.003	.002	.210	M (1960)	-.013	.026	-.091
I				I			
Investment dependence 1959-61 average	-.009	.004	-.374*	Investment dependence 1959-61 average	-.058	.045	-.231
		R=0.426				R=0.250	
Constant -0.069		R ² =0.181		Constant 1.512		R ² =0.062	
N=32		S.E.=0.095		N=32		S.E.=1.122	

Y_t = Percent change in percent of labor force in agriculture between 1960 and 1970
 M = Percent of GDP in mining 1960
 I = Investment dependence, balance on nonmonetary sectors capital, 1959-61
 * $p < .05$

Y_t = Percent change in KWH electricity consumption per capita between 1960 and 1970
 M = Percent of GDP in mining 1960
 I = Investment dependence, balance on nonmonetary sectors capital, 1959-61

findings for neither equation 4 nor equation 5 support Chase-Dunn's asserted significant negative effect of dependence on economic development.

Conclusion

Dependency theorists argue that foreign capital inflows and the vested interests of foreign nationals associated with such financial movements impede social and economic development by interfering with balanced, independent development. Neoclassical economics and modernization theory contend the opposite: that foreign capital inflows and associated technological transfers should act as engines of growth. Chase-Dunn's empirical work supports the first hypothesis. However, his choice of both dependency variable and theoretical model are questionable. The use of the balance of payments summary item, balance on nonmonetary sectors capital, more precisely fits the theoretical description of dependence: penetration of an economy by foreign capital. Cross-sectional analysis using this alternative variable indicates that investment dependence has either no significant effect or a significant positive effect on various measures of economic development. In no instance do we find evidence for the dependency theory assertion that investment dependence has a significant retarding effect on economic development.

Our findings support Delacroix (1977) who has criticized world-system and dependency theories from a somewhat different angle. De-

lacroix compared the effects of two independent variables on 1970 GNP/capita. The first of these was an external dependency variable, the extent of a country's specialization in raw materials export. The second and intrasocietal variable, secondary school enrollment, was conceptualized as the information-processing capacity of the society. Delacroix found in his regression analysis that, with 1955 GNP/capita controlled for, the dependency variable had no significant effect on 1970 GNP/capita. However, the school enrollment variable did have a significant effect in one of the two versions of the regression equation.

In the days when the effects of foreign dependency on development in the less-developed nations were discussed on an anecdotal or case historical basis, there was little agreement in the literature as to whether these effects were generally positive or negative. It seems clear that although this issue now is subjected increasingly to more powerful methodological strategies such as panel multiple regression analysis, the controversy is likely to continue for some time. At the present time, however, it appears that the first wave of enthusiasm for dependency theory is coming under attack. We would insist that the evidence does not support dependency theory's assertion that foreign capital impedes economic development in the nations of Latin America, Asia and Africa.

Mark Stumpp
Robert M. Marsh
Deborah Lake
Brown University

APPENDIX

Sample of Developing Nations Used in Equations 1 and 2*

Argentina	Indonesia
Bolivia	Iran
Brazil	Iraq
Burma	South Korea
Ceylon	Malaysia
Chile	Morocco
Republic of China	Nicaragua
Columbia	Nigeria
Cambodia	Pakistan
Dominican Republic	Paraguay
Ecuador	Peru
El Salvadore	Philippines
Ethiopia	Sudan
Ghana	Surinam
Guatemala	Thailand
Haiti	Tunisia
Honduras	Turkey
India	

*Our data are available for these 35 countries because they refer to 1959-61 and 1960. By using 1950 data, Chase-Dunn (see Table 1, equation 1 above) has an N of only 25 countries.

REFERENCES

- Chase-Dunn, Christopher
1975 "The effects of international economic dependence on development and inequality: a cross-national study." *American Sociological Review* 40:720-38.
- Delacroix, Jacques
1977 "The export of raw materials and economic growth: a cross-national study." *American Sociological Review* 42:795-808.
- Frank, Andre Gunder
1969 *Latin America: Underdevelopment or Revolution*. New York: Monthly Review Press.
- International Monetary Fund
1959- Balance of Payments Yearbook. Washington, D.C.: International Monetary Fund.
- International Bank for Reconstruction and Development (World Bank)
1972 *World Tables*. Washington, D.C.: IBRD.
1973 *World Bank Atlas*. Washington, D.C.: IBRD.
1976 *World Tables*. Washington, D.C.: IBRD.
- Organization for Economic Cooperation and Development
1972 *Stock of Private Direct Investments by D.A.C. Countries in Developing Countries: End 1967*. Paris: OECD.

REPLY TO STUMPP, MARSH,
AND LAKE

The comment by Stumpp, Marsh and Lake on my research on dependence effects is a wel-

come event. Except for a research note by Kenneth Bollen (1977), I have not received any critical responses to what I considered to be a most controversial finding. I have, however, continued to work on the problem of dependence effects with Richard Robinson and Volker Bornschier, and a number of our new findings are germane to the problems raised by Stumpp et al. They correctly criticize my measure of investment dependence, but their own solution is unsatisfactory. Instead of measuring the total accumulated stocks of foreign investment, they measure the recent inflows of new foreign capital. They acknowledge that this does not constitute a proxy for the total value of foreign capital in a country but argue that this is justified theoretically because economists usually study the effects of flows on other flows. Thus they reformulate the theoretical problem and ask what are the effects of foreign capital inflows on development? This reformulation precisely captures the difference between dependency theory, which studies development from a structural and institutional perspective, and most development economics, which tends to ignore such factors.

I was attempting to measure the extent to which a national economy is penetrated by foreign capital. This approach suggests that foreign ownership constitutes a form of control which is part of the structure of domination by which the core of the world economy exploits the periphery. On the other hand, most development economists see foreign capital as a resource which flows into a country causing economic growth. These two different theoretical approaches often lead to the use of different measures.¹ Researchers sympathetic to dependency theory prefer stock measures to estimate the overall structural dependence of a national economy. Researchers who prefer to ignore the hierarchical aspects of the world economy more often use flow measures.

Robinson, Bornschier and I have undertaken a systematic comparison of sixteen cross-national studies of the effects of investment dependence on economic growth (Bornschier et al., 1978). We try to account for the discrepant results which these studies display. One of the main differences between the studies, which accounts for different results, is whether they employ measures of stocks or measures of flows. Those using flow measures tend to find positive effects on growth; while those using

¹ Stock measures are estimates of the total value of foreign direct investment in a country at a particular point in time. Flow measures are estimates of recent inflows of foreign capital. The two measures usually are uncorrelated.

measures of stocks find negative effects. Following the lead of Colin Stoneman (1975), we surmise that this may be a substantive finding. We test this in an analysis which includes both stocks and flows in the same equation.² The results confirm that the stock measure of overall dependence on foreign capital has negative effects on growth, while the measure of recent inflows of foreign investment has positive effects.

This finding is consistent with earlier results (Chase-Dunn, 1975; Bornschier, 1975) which showed the importance of time lags for the effects of stocks of foreign capital. Short-run estimates (five years) often show small positive effects; while longer-run estimates (ten to twenty years) show large and significant negative effects.

Thus we have found that flow measures have positive effects, while stock measures have negative effects which increase with longer lags. These results tend to confirm the hypothesis that current inflows of foreign investment capital cause short-term increases in growth due to the contribution to capital formation and demand as foreign corporations purchase land, labor and materials, and start up production;³ while the long-run structural distortions of the national economy produced by dependence on foreign investment and the exporting of profits produces slower rates of national growth over time. In this light, the ambivalence of policymakers in peripheral countries toward foreign capital becomes more understandable.

Christopher Chase-Dunn
Johns Hopkins University

REFERENCES

- Bollen, Kenneth
1977 "Dependency, population and development." Unpublished manuscript, Department of Sociology, Brown University.
- Bornschier, Volker
1975 "Abhaengige industrialisierung und einkommensentwicklung." *Schweizerische Zeitschrift fuer Soziologie* 1:67-105.
- Bornschier, Volker, Christopher Chase-Dunn and Richard Robinson
Forth- "Cross-national evidence of the effects of coming foreign investment and aid on economic

growth and inequality: a survey of findings and a reanalysis." *American Journal of Sociology*.

Chase-Dunn, Christopher

1975 *International Economic Dependence in the World-System*. Ph.D. dissertation, Department of Sociology, Stanford University.

OECD (Organization for Economic Cooperation and Development)

1972 *Stock of Private Direct Investment by DAC countries in Developing Countries: 1967*. Paris: OECD.

Stoneman, Colin

1975 "Foreign capital and economic growth." *World Development* 3:11-26.

REPLY TO WALSH

In his critique of our analysis of recent farm worker movements in the U.S. (Jenkins and Perrow, 1977), Walsh (1978) charges that I adhere to a monocausal hypothesis of environmental determinism in explaining the outcome of movement challenges. This position in turn leads me to commit a host of sins: ignoring important internal differences between movements launched among farm workers; ignoring the independent role of the "great man in history" (namely, Cesar Chavez); and failing to predict the ability of UFW to survive and eventually prevail in a hostile political environment. Walsh's critique rests upon a number of interpretive errors, both in presenting my argument and diagnosing the experience of farm worker challenges. For the sake of space, my rejoinder will center around the theoretical issues that divide us, though the problems of historical interpretation ultimately possess equal importance (for a detailed historical account, see Jenkins, forthcoming).

Since the charge of environmental determinism forms the basic thrust of Walsh's critique, I'll begin with it. The fact that the terms are Walsh's and not mine suggests the strawman character of his critique. In making the charge, Walsh contends that I reject a more plausible counterhypothesis, which he labels a "reciprocity of influence cycle between a movement and its environment" (p. 110). In the process of making this charge and advancing his counterhypothesis, several important elements of my argument get lost. First, what sense does one make of an exploited, oppressed population like farm workers that remains acquiescent despite these enormous grievances and then suddenly mobilizes to challenge this state of affairs *only* when professional organizers, funded and encouraged by established liberal, religious and labor orga-

² The measure of stocks we use is much better than the one I employed in my earlier work. We use the OECD (1972) estimate of the total book value of stocks of foreign direct investment.

³ It is somewhat puzzling that Stumpp et al. do not find statistically significant positive effects of flows in their analysis. Perhaps this is because the ten-year time lag they employ is longer than the short-term positive effects of inflows of foreign capital.

nizations, are thrust into the scene? I take this as indicating that farm workers (unlike middle-class women, professionals or skilled blue-collar workers) lacked the indigenous resources and in-group ties to mobilize "spontaneously" without the insertion of this outside stimulus. (Several have misread our argument to apply to all movements; they fail to note its specific restriction to powerless groups, cf. Fendrich, 1977). Walsh remains mute on this issue, but it is central to the basic argument. The implication is that without processes in the political environment of an aggrieved, powerless population that gives rise to the interjection of movement leaders possessing organization skills and having access to a minimum of resources, no movement organization will form. (For a collaboration of my position see the recent piece by Lawson, 1978, on the tenant movement in New York City.)

Second, what weighting does one assign to the internal characteristics of movements (e.g., organization, leadership, tactical choices) relative to environmental conditions in determining movement outcomes? Ultimately the question resolves into one of power. Does the movement have *direct* control over the resources needed to bring about changes or does it depend upon *outside* support? My argument was that changes in national political alignments make it possible (though not necessary) for a protest strategy (boycotts) to succeed. By its nature, protest is an outside power strategy in that it seeks to apply leverage by mobilizing third-party constituencies (consumers) and organizations (the Catholic church, the AFL-CIO) who then in turn pressure movement opponents (growers) to make changes. Movements of the powerless depend upon protest for their success in forcing targets to make or accept changes. Both farm worker movements tried boycotts; only the UFW grape boycott succeeded. Walsh is clearly wrong to claim that the NFLU's boycott attempt was not serious (p. 110), as both Galarza's (1970;1977) account and my own interviews with the NFLU leadership amply prove. The NFLU was abandoned by its few liberal and labor allies and illegally forced by the National Labor Relations Board and the California courts to halt the boycott. By contrast, UFW won every one of its contracts up until the legalization of collective bargaining on the basis of widely supported boycotts. Moreover, as Walsh himself admits, the organizational resources on which UFW operated derived primarily from outsiders (from students and radical clerics staffing the boycott offices and union hiring halls to radical lawyers donating legal assistance), not from indigenous farm worker resources.

Though money is not the all-critical resource that some resource mobilization theorists have argued (cf. McCarthy and Zald, 1973; 1977), it is indicative that less than one-third of UFW's budget for the first ten years of its operations came from member dues. The other two-thirds came from outside contributions. By contrast, the NFLU had few such resources, either full-time volunteers (which could be counted on one hand) or financial contributions. This does not mean that movement organizations lack direct control over the resources needed to bring changes. Those that succeed depend upon outside support; those that don't, fail either because they do not seek support through protests (which I argued was the case for the AWOC) or because the audiences at which protest actions are aimed fail to respond sufficiently to make a difference. The likelihood of such a supportive environment depends in turn on national political alignments (the posture of elites, the political influence of the liberal-labor coalition), not simply on the persuasiveness and charisma of movement leaders.

The key point, then, is that my argument places the dynamics of movement-environment interaction within a macropolitical context and that the interaction between movements and their supporters are characterized as power relations. It was a qualitative shift in the pattern of national political alignments that made possible successful insurgency by the powerless, though success was not environmentally guaranteed. Such movements could fail tactically to secure support; or they could adopt an ideology that scared off liberal support (e.g., the Black Panthers). But there was a world of difference between the political period of the NFLU, bounded as it was by the purges of the industrial unions beginning in 1946 and the peak of McCarthyist hysteria in 1954, and that of the late 1960s. Most important of all, given the resource dependence of movement organizations of the powerless upon outside supporters, Walsh seriously distorts reality by calling this relationship one of reciprocal influence. Walsh's term suggests pluralistic bargaining, a far cry from being able to manipulate an already favorably-inclined coalition of organized labor and liberals. The chief problem is not noticing Chavez's cleverness (which is indisputable) but analyzing why this support coalition held out against incredible odds so long.

Walsh raises a number of other issues which need to be touched upon. Walsh leaves unclear what he means by the "great man in history" argument. If charisma is meant, then anyone who has seen Chavez in action will agree that

he lacks it. What gives Chavez persuasiveness is dogged personal commitment to a just cause, not claims to divine insight. If what is meant is the role of the individual in the making of history, then he has an issue, but not one to which there is a sociological answer. I don't think my analysis rejects the notion that Chavez played an important role but defines the conditions circumscribing effective action. As for the difference in the ethnic base of the two unions, Walsh correctly notes the basic divergence. However, his own conclusions on this lend weight to our environmental hypothesis. As Walsh notes, the importance of this different social base was not that it lead to successful strikes for the UFW, but that it allowed "Chavez and his cadre to encourage critical support from the post-Vatican II church" (p. 110, italics added). As argued in my forthcoming book, this ethnic difference did not provide union contracts. In other words, the important issue was the symbolic significance that the minority status of farm workers assumed in the eyes of outsiders.

As for the post-1972 period, there are several points to be made. First, we were not engaging in futurology but in historical explanation. Moreover, Walsh's own analysis highlights the critical role played by the political environment: "The critical turning point in the UFW's struggle for survival came with the election victory of Jerry Brown" (p. 111). Though UFW played an independent role in electing Brown, this was *after* the union had already proven itself a partial success and had thereby managed to acquire an independent resource base. Significantly it was not farm worker votes that UFW mobilized to support Brown but urban Chicanos and liberals.

J. Craig Jenkins

University of Missouri, Columbia

REFERENCES

- Fendrich, James M.
1977 "Unions help faculty who help themselves: a partisan view of a collective bargaining campaign." *The American Sociologist* 12:162.
- Galarza, Ernesto
1970 *Spiders in the House and Workers in the Field*. Notre Dame: University of Notre Dame Press.
- 1977 *Farm Workers and Agribusiness in California*. Notre Dame: University of Notre Dame Press.
- Jenkins, J. Craig
Forth- Movement and Power: Farm Worker Incoming insurgency in the U.S.
- Jenkins, J. Craig and Charles Perrow
1977 "Insurgency of the powerless: farm worker

movements in the U.S. (1946-1972)." *American Sociological Review* 42:249-68.

- Lawson, Ron
1978 "Who pays the piper: the effects of outside funding on social movements of the poor." Unpublished paper, Department of Urban Studies, Queens College, City University of New York.
- McCarthy, John and Mayer Zald
1973 *The Trend of Social Movements in America: Professionalization and Resource Mobilization*. Morristown: General Learning Press.
- 1977 "Resource mobilization and social movements." *American Journal of Sociology* 82:1212-41.
- Walsh, Edward
1978 "On the interaction between a movement and its environment (comment on Jenkins and Perrow, ASR April, 1977)." *American Sociological Review* 43:110-2.

MILITARY INEQUALITY: REPLY TO HAUSER*

In a recent paper I examined the universalistic standards argument which has been utilized to explain black inequality in the military (Butler, 1976a). Taking promotion time as a measure of inequality, the mean months in service for black and white army enlisted men to make grades E4, E5-E6, and E7-E9 were compared. After controlling for the AFQT (Armed Forces Qualifications Test), education and military occupational type (Technical or Nontechnical), the following was concluded:

... This paper presents data which suggest that universalistic criteria are not sufficient to explain black inequality vis-à-vis promotion time. When black and white enlisted men are matched on key universalistic criteria (civilian education, Armed Forces Qualifications Test, and occupational type), blacks consistently take more time to make grade than whites. The findings suggest the salience of the notion of ascription: when ascription is based on race, it becomes a component part of racism. Thus, one unavoidably is led by the data ... to conclude that the black enlisted man is subject to inequality, which is not the result of failure to meet universalistic criteria, but, rather, the result of the racist actions of real life people. (p. 807)

In a recent comment on that article, Hauser (1978) questioned the reliability of the conclusions. He noted:

Unfortunately, differentials in mean months to promotion bear no necessary relationship to But-

* I would like to thank Frank Bean for comments on this paper. The author accepts full responsibility for paper content.

ler's conclusion and for this reason can neither confirm nor disconfirm them. The problem is that there is no closure in the military population, either with respect to movement out of the military into the civilian population (or death) or with respect to movement between ranks. There is no self-evident interpretation of the temporal differences between black and white rank incumbents in months in service to make current rank because persons occupying a given rank do not represent the experiences of all persons who have been promoted to that rank, nor of all persons exposed to the "risk" of promotion. Some persons who were promoted to the given rank will subsequently have left the military, while others will have been promoted to higher ranks. Moreover, months in service to make current rank is not a valid inverse measure of the likelihood of promotion to that rank because it does not take account of the experiences of persons who might have been promoted to that rank but were not. (pp. 115-6)

I certainly appreciate Hauser's careful attention to my paper; comments which generate additional thought are always needed and welcomed. There are a number of places, however, where I find his arguments confusing, others where I disagree, and finally others which reflect an unfamiliarity with the literature on race and the military. For example, commenting on the negative relationship between promotion and AFQT Hauser notes "that the current incumbents of a given rank would include a disproportionate number of those men who took longest to achieve it" (p. 116). This means, he adds, that

higher scoring incumbents of a given rank would include a relatively larger number of men who had been passed over for promotion than the lowest scoring incumbents of the same rank because the most promising of the higher scoring incumbents would already have been promoted. (p. 116)

Hauser says this "alone might account for the negative relationship between AFQT scores and mean time to promotion to current rank" (p. 116). This hardly seems surprising if any relationship exists at all between ability and rate of promotion. More important, it is not at all clear that these remarks pertain to the issue at hand: racial differences in promotion time. Unless the processes he is talking about can be shown to differ by race (which demonstration Hauser does not make), they are irrelevant to the point I was making.

One of the excellent points made by Hauser has to do with the military not being a closed population. In his words, "there is no closure in the military population, either with respect to movement out of the military into civilian population (or death) or with respect to movement between ranks" (p. 115). It is important to note here that the data utilized in my original

article were taken from the military active file for the year 1973. This means that my study population consisted only of those personnel on active duty. Put differently, I did not examine personnel who had left the military. Hauser's basic argument is that the findings cannot be confirmed (or disconfirmed) because some of those persons who were promoted already might have departed the military. Although this point might appear to be well taken, the issue of persons moving out of the military, together with the question of whether departure explains differences in mean promotion time for blacks and whites, already has been aired and researched by scholars of military institutions. The results were published and made available to the public and scholarly community a year following my original paper and are discussed below (Department of the Army, 1977).

In order to address Hauser's point data are needed over time. We need to look not only at the active file (in my data 1973), but also at the active files of years past, which of course are now called inactive files. Inactive files thus consist of two kinds of people, some of whom have departed the military, and some of whom are still in the military. Of course, if a person is on the inactive file of, say, 1971 but has not left the military, he is also on the active file of 1973. Put simply, inactive files allow us to view the military as it was in a given year and therefore are inclusive of persons who have since left the military.

Table 1 presents the analysis of inactive files for mean months in service to make current grade for the years 1971 and 1972. Grades are not grouped as in my original article, but the patterns remain the same as revealed in my 1973 data. In 1971 and 1972 blacks consistently took more time to make their current grade. When control variables are introduced, the same patterns persist, with some of the same anomalies discussed in my original article. As a matter of fact, when difference scores were calculated for black and white mean months over time the difference was greatest in earlier years. This means that although blacks have made some gains, the obvious differences remain.

Another alternative hypothesis which Hauser introduces to account for differences by race in mean promotion time is "the fact that whites have better opportunities than blacks for civilian employment" (p. 116). From this Hauser speculates that "relatively more of the less meritorious whites than of the less meritorious blacks will leave the army, and this differential in retention could account for the racial differences in mean time to promotion

Table 1. Mean Months in Service to Make Present Grade by Race and Year (1971 and 1972) *

Grade	Mean Months in Service 1971	
	Black	White
E4	14.50 (22,538)	12.25 (157,983)
E5	37.66 (11,129)	27.10 (80,542)
E6	102.20 (11,310)	95.44 (35,990)
E7	171.45 (6,649)	157.75 (26,667)
E8	216.85 (1,312)	206.74 (8,177)
E9	254.33 (174)	243.37 (2,124)
Grade	1972	
	Black	White
E4	15.60 (18,479)	12.74 (112,871)
E5	40.78 (12,809)	33.08 (63,517)
E6	97.79 (14,702)	94.30 (43,447)
E7	163.64 (8,212)	151.85 (31,093)
E8	220.72 (1,602)	206.77 (8,741)
E9	258.41 (292)	244.28 (2,805)

* N in parenthesis: Source: Department of the Army (1977).

among current incumbents" (p. 116). While seemingly compelling, this argument must be recognized for the speculation it is. First of all, one suspects that relatively more of the more meritorious whites than of the more meritorious blacks will *also* leave the army. Indeed, if Hauser had taken more time with the literature on race and the military, he would not even have made this speculation. What the literature shows is that with the advent of the all-volunteer service, more "meritorious" whites *have* left the military and indeed are not joining the military in the first place. Because of this an unprecedented trend has developed in the military. Using educational levels as a measure of merit, Moskos (1978:11) notes:

Contrary to national patterns, however, the intersect of race and education is quite different among male entrants in the all-volunteer army. Since the

end of the draft, the proportion of black high school graduates entering the army has exceeded that of whites, and this is a trend that is becoming more pronounced. [The trend has reached the point where] high school graduates accounted for 60.5 percent of entering blacks compared to 40 percent of entering whites. In point of fact, today's army enlisted ranks is the only major arena in American society where black educational levels surpass that of whites, and by a significant margin!

Education is a good measure of merit because it predicts performance and other variables like promotion rates better than any variable utilized in the study of military institutions (Moskos, 1978; Butler, 1976a). In essence, however, the issue is not whether meritorious whites leave the military faster than meritorious blacks (which could happen simply because whites leave faster than blacks in general), but rather whether departure rates vary by ability (as measured by variables other than education such as the AFQT) and race. Hauser does not bother to utilize the "literature as data" to make his point (and I believe that the existing literature will not confirm his argument); put another way, Hauser does not show that departure rates vary by ability, and thus his argument is simple speculation.

Hauser makes another comment which the literature on race and the military has already aired. To make his main point, he constructs a table from my data showing the percentage distribution of military rank by race, AFQT score, education, and military occupation.

Based on these figures he notes:

Within levels of schooling, AFQT, and military occupation, black enlisted men without exception are less likely than whites to be in E4 relative to E5 through E9. . . . Black men were also *more* likely than whites to be grades E7 to E9 relative to the numbers in E4 through E6. . . . Most noteworthy is the consistent finding that the high AFQT men are less likely to be in rank E4 and more likely to be in ranks E7 through E9 than men of low AFQT scores with the same race, education and military occupation. . . . Most important, conditional on promotion to rank E4, Butler's data show that blacks are more likely than whites of similar AFQT, schooling and military occupation to be in the higher enlisted ranks, and the differentials favoring blacks in the chance of occupying a high rank are larger where the apparently opposite tendencies in mean time to promotion are greater. (pp. 116-7)

Hauser then asks, "Do these data suggest a sociological interpretation of army promotional practices in the enlisted ranks which is opposite from that made by Butler? That is, are blacks favored over whites by army promotional practices?" (p. 118). Anyone familiar with the literature would give an unconditional *no* to Hauser's questions. The reason blacks are in higher ranks has nothing to do with

blacks being favored over whites. It has to do with the very high reenlistment rate of blacks in the army. That is, blacks are in high ranks because more of them stay in longer than whites. This is also one of the issues which has generated so much debate about the all-volunteer force (Janowitz and Moskos, 1974; Schexnider and Butler, 1976). The high reenlistment rate, which has been noted before in the literature (Butler, 1976b), is brought up to date in a recent paper. Moskos (1978) notes that

the proportion of black noncoms can be expected to increase further owing to the higher than average re-enlistment rate of black soldiers. In FY 1977, re-enlistments at the end of the first term were 48.5 percent for blacks and 30.6 percent for whites. . . .

Certainly one should continue to see blacks at the higher levels of military rank due to their high reenlistment rate.

Among other questions Hauser asks,

Does a control for military occupation adequately reflect differentials in promotional opportunities between branches of service? How are differentials in the size and composition of entering cohorts of enlisted men and of their chances for promotion affected by the differing conditions of the pre-Vietnam and Vietnam war? (p. 116)

It was clear that my data pertained to the army, not the navy, air force or marines. Also, since we know that blacks fought in Vietnam in higher proportions than their percentages in the civilian arena and the army, and that war time service is highly correlated with faster promotions, then if there is a "Vietnam effect," it is plausible to say that blacks should have benefitted more. In short, most of Hauser's comments are mere speculation which often seem not to consider the mass of literature on race and the military. (The best point made by him is that the data do not allow us to look at promotion between ranks, it only allows us to examine promotion from the time a person entered the service. The latter was my objective and the former should be examined in the future.) And what is wrong with speculation? Nothing, when presented for what it is. Hauser, however, makes it sound as if his alternative interpretations of the data are rooted in fact, when they are not. But to the extent that this exchange may serve to stimulate us all to seek new data and facts to test our ideas, all work in this area will benefit tremendously.

In closing, although Hauser "cannot accept Butler's interpretation of his data," I think that the material presented here reinforces the original findings and theoretical interpretations.

John Sibley Butler
University of Texas, Austin

REFERENCES

- Butler, John Sibley
1976a "Inequality in the military: an examination of promotion time for black and white enlisted men." *American Sociological Review* 41:807-18.
1976b "Assessing black enlisted participation in the army." *Social Problems* 23:558-66.
Department of the Army
1977 *Measuring Changes in Institutional Racial Discrimination in the Army*. Pamphlet No. 600-43. Headquarters, Washington, D.C.
Hauser, Robert
1978 "On inequality in the military (comment on Butler, ASR October, 1976)." *American Sociological Review* 43:115-8.
Janowitz, Morris and Charles C. Moskos, Jr.
1974 "Racial composition in the all-volunteer force." *Armed Forces and Society* 1:109-24.
Moskos, Charles C., Jr.
1978 "The enlisted ranks in the all-volunteer army." In John Keeley (ed.), *The Military in American Society*. Charlottesville: University of Virginia Press. In Press.
Schexnider, Alvin J. and John Sibley Butler
1976 "Racial composition in the all-volunteer system: a reply to Janowitz and Moskos." *Armed Forces and Society* 2:421-32.

REPLY TO RONALD L. SIMONS

In his "The Meaning of the IQ-Delinquency Relationship," Ronald L. Simons (1978) suggests the possibility that our "Intelligence and Delinquency: A Revisionist Review" (Hirschi and Hindelang, 1977) may not be the last word on the subject. He says we are "blatantly wrong" in our interpretation of IQ scores, that we "contribute little to the topic other than obfuscation," that we "ignore a large body of literature" and consequently present an "outdated" and "naive" view quite at odds with the "contemporary" understanding of "experts" and "researchers" in the area. Professor Simons goes on to point out that the simple causal model we present must compete with the "several models" he could supply. He suggests that all these models are equally plausible; if so, the chances are very good that he is right and we are wrong. As if this were not enough, Professor Simons draws out some of the implications of our point of view: "accepting Hirschi and Hindelang's contentions concerning IQ tests would demand that sociologists view blacks as less intelligent than whites" (p. 268); "Hirschi and Hindelang would be forced to draw the conclusion that college students are becoming less intelligent" (p. 269).

The decision to attempt to reply to Professor Simons's comment was not easy to make. He obviously is blessed with an inventive mind, and he appeared to be an expert in the area of intelligence (which we are not). There was the unalterable fact that his comment was scheduled for publication in the *American Sociological Review*. (Comments submitted to this journal do not automatically appear in print; they are rigorously screened by leading authorities in relevant fields.) Then, too, Professor Simons's critique suggests that our ignorance is so glaring and our mistakes so numerous that a persuasive reply would be beyond our capacity. Against all this, we were able to come up with only two minor virtues: we can read and we can (sometimes) predict the future.

The ability to read presumably is acquired. We recommend it to Professor Simons. It is fun and it is often useful in scholarly pursuits. For example, using this acquired skill Hindelang was able in five minutes to deduce from Professor Simons's critique that Professor Simons had not read the primary material on which his comment is based. Using the same skill in the library, Hirschi in a matter of minutes confirmed the Hindelang hypothesis. A brief phone call to Professor Simons, and this virtually certain hypothesis was established fact.

Let us now demonstrate our clairvoyance. In 1975 we anticipated Professor Simons's 1978 attack on our paper, almost to a word. (Modesty requires a confession: in 1976, the editor of the *American Sociological Review* also predicted Simons's comment with amazing accuracy.¹) Seeing the future is one thing; doing something about it is another. What did we do about it? Well, we decided we didn't want to write this reply to Simons, so we wrote a paper that would leave him with nothing to say. Unfortunately, we neglected to consider Professor Simons's attitude toward reading. Being careless of words, Professor Simons simply supplied those we left out, and then responded to a paper we didn't write.

Given this background, it seems pointless to try to explain the logic of our paper to Professor Simons. Still, if he is right in his concep-

tion of IQ, there is always hope. In our paper, IQ is treated as an *exogenous* (e-x-o-g-e-n-o-u-s) variable. The sources of IQ differences therefore are not our concern. It matters not one iota to our argument whether IQ is transmitted through the genes, the environment, the U.S. Mail (via a postcard from Rosenthal and Jacobson), or some combination of these. All our argument requires is that IQ differences be present prior to school entrance and that these *differences* remain reasonably stable over time. (These conditions are met in all data we know anything about).

And that's about it.

We believed when we wrote the paper, and we believe now, that no sociologist should be *required* to take a stand on the heritability issue when that issue is irrelevant to the question at hand. Our paper therefore ignores (*pointedly* ignores) heritability. Professor Simons still manages to infer that ours is a neogenetic theory that would "demand" racist conclusions by other sociologists. Since he does not tell us how he reached this convenient conclusion, we can only assume he made it up to bedazzle ASR referees.

Had we discussed the heritability question it would have been impossible to be as naive as Simons imagines. We could not have maintained that heritability of IQ makes observed scores "independent of cultural or social influence" any more than we could argue that the heritability of characteristics of corn makes it impervious to the effects of water, temperature, and sunlight. Moreover, it is hard to imagine how we could be forced to assume that stunted corn is necessarily genetically inferior to well-tended corn.

Our *necessary* assumptions about IQ scores are shared by the IQ-change studies extensively cited by Simons as evidence of our stupidity. In fact, research in the IQ-change area typically assumes that important differences are established much earlier than our argument requires—i.e., in the first few years of life (e.g., Whimbey and Whimbey, 1975:41). Some stability of the results is also automatically assumed (otherwise, why spend such lavish sums and why crow so loudly about the results?) There would seem then to be no disagreement of substance between the contemporary view and our own.

Incredibly, Professor Simons appears to agree with his own experts (and with us) that the test-retest correlations of IQ tests administered many years apart are strong. This is what is usually meant by the word *stable*. But whenever the urge strikes him, Professor Simons grandly assumes "that IQ scores are unstable" (p. 268). Which is it going to be? If the

¹ "The immediate reaction will be—and again I mean nonsensical comments submitted to ASR for publication—that intelligence is socially constructed, that it is not unidimensional at all, that it depends on motivation as well as intellectual function. . . . It should be a little more explicit in the paper that *it does not matter* what attitude one has to the meaning of the IQ test, so far as the paper's argument goes." Morris Zelditch, Jr., personal communication, December, 1976. Italics in original. Quoted with permission.

differences are stable over time, and Simons agrees they are stable over time, then how can our assumption that they are stable over time be evidence that *we* don't know what we are talking about? If Simons agrees the scores are stable over time, how can he assume they are unstable over time? These little conundrums are easily solved. They are in fact the key to an understanding of the differences between Simons's position and ours: our discussion applies to the world that is; his only on occasion applies to this world. More often it applies to fantasy worlds that could or should be. Unfortunately, Professor Simons doesn't know when he is talking about one of his fantasy worlds and when he is talking about this one. How otherwise explain his delightful suggestions that IQ scores at school entrance may be caused by delinquent acts committed at age 15, or that negative attitudes developed before IQ differences begin to emerge (ages two to four) may be the cause of both low IQ scores and delinquency?

Let's get back to the connection between the world that is and worlds that could be. In our article, we summarized criticisms of Rosenthal and Jacobson's famous *Pygmalion in the Classroom*; we pointed out that nine attempts to replicate the results of this study had failed. (A fact that Professor Simons ignores in questioning the adequacy of our critique). We frankly weren't certain what bearing the Pygmalion results had on our argument, but we didn't believe them² and we knew they were somehow considered contrary evidence, so we took them on. Now Professor Simons throws against us five studies reporting the results of efforts to "increase the IQ scores of low social class and minority group individuals" (p. 269). We are still uncertain what bearing all this has on our argument. In fact, we could easily say we are delighted with the success of such efforts. According to our theory, these successes should produce important reductions in the delinquency rate among those receiving the treatment. We could note, too, that according to Simons's models these efforts are pointless exercises which neglect the fact that IQ scores are unstable, that delinquency comes before IQ, and so on. But that would be too easy. Besides, we don't believe some of the results Simons reports, and we consider his treatment of the IQ-change literature irresponsible.

The "large body of literature" cited by Si-

mons could not be obtained in the two weeks granted us to prepare this reply. Fortunately, however, we were able to find the sources Simons admitted relying on for his comment: Whimbey and Whimbey (1975) and Ryan (1976). In addition to the items cited in our original paper, Simons cites thirteen other works. Of these, eleven can be found in Whimbey and Whimbey (1975) or Ryan (1976). There is an uncanny resemblance between Simons's summaries and the summaries in these sources. Compare for example:

Ryan (1976: 55-7)

I am referring to a study on the Harlem schools, designed and conducted by James Jones, Kenneth Clark and the staff of the Haryou Planning Project. I replicated the study in schools in Boston and achieved almost identical results. . . . In early grades—second, third and fourth—there are no significant differences between the presumably culturally deprived children and others. . . . In the fifth and sixth grades . . . the culturally deprived children began to fall behind. And by the eighth grade the differences are large and clear-cut. . . . Clearly, whatever mechanism accounts for these strange findings is found primarily in the interaction of the school with the child.

Simons (1978: 269)

James Jones (1965), Kenneth Clark, and the staff of the Haryou Planning Project studied two kinds of schools—middle-class black and lower-class black schools. They found that in the early grades (second, third, and fourth) there was no significant difference in reading comprehension scores between the children in the two kinds of schools. In the fifth and sixth grades, the children from the lower class schools began to fall behind. By the eighth grade there were large differences between the two groups. These findings, which were replicated by Ryan (1965) . . . suggest [that] something about their interaction with the school system seems to stagnate their growth.

What Simons asserts to be Ryan's (1965) replication of the Haryou Planning Project results turns out to be a replication of Ryan's (1976:55) assertion that his study "achieved almost identical results"; Ryan's (1965) "replication" is an unpublished background paper. Simons's *failure to cite* Ryan's (1976) summary leaves the reader with the clear impression that Simons has seen both the original Haryou study and Ryan's "replication," when, in fact, he has seen neither. As we shall show, Simons's dependence on Whimbey and Whimbey (1975), which he does list among his references, is even more damaging to his credibility as a serious scholar.

Whimbey and Whimbey are described by Simons as researchers in the area of intelli-

² Professor Simons will undoubtedly take this statement as a confession of genetic bias. We hate to disappoint him, but the statement reflects only a bias against labeling theory (a bias shared by the data) not against environmental theories in general.

gence. They provide "the contemporary view of IQ tests" he quotes at some length. It turns out that the bulk of Simons's remaining citations and summaries of the data—even the published materials—are in Whimbey and Whimbey.

Whimbey and Whimbey (1975:42) describe the Milwaukee project (Heber et al., 1972) as "the most lengthy and comprehensive study ever undertaken on training intelligence." (Little wonder Simons does not hesitate to report the results of such research without feeling it necessary to read it.) In *Reader's Digest* fashion, Whimbey and Whimbey (1975:47) merely allude to an article by someone named Page (1972), who is said to have "outlined what he considered to be methodological complications in the Milwaukee project." Whimbey and Whimbey (1975:47) say that Page's points are "technical and detailed, so they will not be discussed here." Simons should have read Page.

According to Simons's (1978:269) summary of Whimbey and Whimbey's summary—which they tell us is based "partly on a review article by S. Strickland" (Whimbey and Whimbey, 1975:42)—of the unpublished results of the Heber study, Heber and his colleagues". . . found that their Milwaukee project resulted in a mean IQ of 124 for the experimental group, and 94 for the control group."

In going back the one step to Whimbey and Whimbey, we were surprised to find that Simons had shown unwonted generosity in his comment on our paper. He failed to mention that the mothers of the subjects had IQs of less than 75 and that IQs registered by experimental children were as high as 135 (Whimbey and Whimbey, 1975:47). One hypothesis explaining Simons's generosity is that he perceived that this was all just too good to be true.

One step further back brought us to Page's technical review, a saga of his attempts to locate the Milwaukee research report and to explain the miracle it described—a miracle previously reported in *Time*. For many months, neither Heber nor the funding agency (Social and Rehabilitation Services of HEW) would answer Page's letters and phone calls about the project. At last a technical report was obtained from the funding agency. Page (1972:10, 15–6) discovered: (1) assignment to experimental and control groups apparently had produced differences sufficiently large to "destroy any reasonable possibility of pre-treatment equality"; (2) there was good reason to believe that subjects had received training on items included in the IQ tests; (3) that the description of the

treatment is "only rhetoric with no operational meaning." Page's (1972:16) conclusion is that "any one of these [defects] would largely invalidate a study. Together they destroy it."

Whether Page's (1972:8) assessment of the Milwaukee project is or is not accurate,³ there can be no doubt about his observation that the Milwaukee results can be "cited in class [and in the *American Sociological Review*] by professors who have not done their academic homework."

Travis Hirschi

Michael J. Hindelang

State University of New York at Albany

REFERENCES

- Heber, R., H. Garber, S. Harrington, C. Hoffman, and C. Falender
1972 Rehabilitation of Families at Risk for Mental Retardation: Progress Report. Rehabilitation Research and Training Center in Mental Retardation, University of Wisconsin, Madison.
- Hirschi, Travis and Michael J. Hindelang
1977 "Intelligence and delinquency: a revisionist review." *American Sociological Review* 42:571–87.
- Jones, James A.
1965 "Education in depressed areas: a research sociologist's views." Paper at the Fourth Work Conference on Curriculum and Teaching in Depressed Areas, Teachers College, Columbia University, New York.
- Page, Ellis
1972 "Miracle in Milwaukee: raising the IQ." *Educational Researcher* 1:8–16.
- Rosenthal, R. and Lenore Jacobson
1968 *Pygmalion in the Classroom*. New York: Holt, Rinehart and Winston.
- Ryan, William
1965 "Racial imbalance, cultural deprivation, and reading achievement." Mimeo. Advisory Committee on Racial Imbalance and Education, Background Paper No. 7, Massachusetts State Board of Education.
- 1976 *Blaming the Victim*. New York: Vintage Books.
- Simons, Ronald L.
1978 "The meaning of the IQ-delinquency relationship." *American Sociological Review* 43:268–70.
- Whimbey, Arthur and Linda Shaw Whimbey
1975 *Intelligence Can Be Taught*. New York: Dutton.

³ A phone call to Page in March, 1978 confirmed that he had not changed his opinion of the adequacy of the Milwaukee research. Another to Garber (the second author of the Milwaukee progress report) confirmed that no direct response to Page had been made.

STATE BOUNDARIES AND THE WORLD-ECONOMY: REPLY TO STACK*

The comment by Stephen Stack (1978) on my article, "The World-Economy and the Distribution of Income within States," concluded that the world-economy model of national income distribution needs to be supplemented by a model of internal political structure. He based this conclusion on an analysis which showed that a measure of internal political structure, democratization, had a statistically significant effect on income inequality, even when a world-economy factor, the volume of exports, is controlled for.

Stack's findings replicate other such findings, including an analysis that Dan Quinlan and I presented in an article in this journal last year (Rubinson and Quinlan, 1977:615). But the main point of our article, developed in both the discussion and the several analyses, was that the causal direction between democratization and income inequality was problematic. The basic issue, we argued, was the understanding of the relationships between class structure and regime forms. This issue is quite complicated, and rather than discuss it here, I will turn to two other issues raised by Stack's comment. I raise these points by way of clarification and not by way of criticizing Stack's analysis; for I fully agree with his point that the form of political regime affects the class and income structures within a country.

A point with which I do not agree, however, is with the conceptual dichotomy he draws between a world-economy model and an internal-structure model. For the fundamental idea in the world-system perspective is that such a distinction does not exist. The world-system is not an *international* perspective in the sense that its contribution is to stress that factors external to state boundaries affect what goes on within those boundaries. Rather, the world-system perspective begins with the idea that what we are accustomed to thinking of as separate systems are simply regional inequalities or differences within a single division of labor, the capitalist world-economy.

The difficulty in understanding this point is itself a consequence of the functioning of this world-economy. For as Immanuel Wallerstein (1974:349) has pointed out, "While, in an empire, the political structure tends to link culture with occupation, in a world-economy the political structure tends to link culture with spatial location." This means that the political pressures in the world-economy tend to produce

local, i.e., national, cultures. As a consequence, our major cultural institution of language reflects this process and institutionalizes the idea that nations or states are separate systems. By this process the cultural system obscures the underlying structural unity of development by creating in our consciousness a fetishism of state boundaries. The point is not that state boundaries are unimportant; they are vitally important. For the effect of the state organization of the world-economy has been to localize, i.e., nationalize, political pressures which are created by the system as a whole. The point is rather that an internal-external dichotomy obscures the fact that state boundaries, political regimes, trade flows, and multinational corporations are all *internal* to the world-economy.

A second issue that Stack's comment raises concerns the concept of state strength, which he correctly points out is a variable that is difficult to measure. But prior to accurate measurement is the problem of conceptualization. State strength refers to the capacity to mobilize resources, both human and financial, through effective decision making. The strength of the state, therefore, cannot be equated with the size of the state apparatus or the degree of repression by the state. Large civil and military bureaucracies are often the result of greater opposition to the state than of the state's strength. The Dutch state of the seventeenth century, the British state of the nineteenth century, and the United States' state of the twentieth century were all relatively small and less autocratic compared with other states of the period. But they were clearly strong states in the sense of being able to mobilize vast resources quite quickly and in creating effective political policies to support the economic interests of their dominant classes (Chase-Dunn and Rubinson, forthcoming).

A comparison of semiperipheral countries in the nineteenth century illustrates well the difference between a strong state and an extensively bureaucratized and autocratic one. For the United States had a very small and limited civil and military bureaucracy; while Prussia had a very large, autocratic, and well disciplined one. Yet both states were very strong in the sense that they were able to effectively take political decisions and mobilize resources to create those policies which enabled them to enter the core of the world-economy. Italy in midcentury had a weak and limited state bureaucracy, while Austria had a large, autocratic, and extensive one. But both states were relatively weak compared with the United States and Prussia in the sense that they were unable to effectively organize their resources

* I would like to thank Chris Chase-Dunn for his help on this comment.

to take advantage of the changing conditions in the world-economy, and consequently they remained semiperipheral states. The reason why the size and character of the state apparatus does not correlate well with state strength arises from the variable conditions that lead to the type of state apparatus that is built by the state-makers: Prussia's extensive civil and military bureaucracy was a response to the defense of its precarious geopolitical position within central Europe; while Austria's extensive and autocratic bureaucracy was built as a response to its disintegrating empire. The United States, in a protected geopolitical position, had no need for an extensive or autocratic bureaucracy; and Italy's largely foreign-controlled territory prevented the establishment of extensive political authority. (See Tardanico, 1978, for a thorough discussion of these issues.)

If the size and character of the state apparatus are not what create strength, then what is the crucial factor? The answer seems to be that state strength follows from the degree to which the interests of an area's dominant economic classes are institutionalized within the state, and the state will be even stronger to the extent that those groups are among the major capital accumulating classes in the world-economy as a whole. This institutionalization of interests leads to a strong state for two reasons. First, such a condition means that the state structure needs to devote relatively little of its resources toward overcoming the hostility of indigenous elite classes; and second, such a condition provides the underlying resource base necessary for effective state action (Tardanico, 1978). Such a coalition of interests was formed in both the United States and Germany between 1840 and 1880; while that same potential coalition was torn apart in Austria and Italy during the same period (Rubinson, forthcoming).

Richard Rubinson
Johns Hopkins University

REFERENCES

- Chase-Dunn, Christopher and Richard Rubinson
Forth- "Toward a structural perspective on the coming world-system." *Politics and Society*.
- Rubinson, Richard
Forth- "Political transformation in Germany and coming the United States." In Barbara Kaplan (ed.), *Social Change in the Capitalist World Economy*. Beverly Hills: Sage.
- Rubinson, Richard and Dan Quinlan
1977 "Democracy and social inequality: a reanalysis." *American Sociological Review* 42:611-23.

Stack, Steven

- 1978 "Internal political organization and the world economy of income inequality (comment on Rubinson, ASR August, 1976)." *American Sociological Review* 43:271-2.

Tardanico, Richard

- 1978 "A structural perspective on state power in the capitalist world-system." Paper presented at the annual meeting of the American Sociological Association, San Francisco.

Wallerstein, Immanuel

- 1974 *The Modern World-System: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century*. New York: Academic.

DIAMONDS ARE TRUMPS: A REPLY TO HOUSE

Professor House's paper in the June issue of this journal is an exception to the general rule that controversy usually arises out of error and misunderstanding. House and I understand one another perfectly well, yet we continue to disagree. Indeed, the burden of his paper was anticipated in my remark (Hope, 1975:342) that "the typical response of sociologists who have accepted the argument of this paper has been to eschew a concept of general status." He is, in fact, ably defending a research tradition which is extensive and well-entrenched, and has achieved the status of normal science in Kuhn's (1970) sense. Why do I want to subvert it?

In my paper the essence of my position was stated thus (p. 341):

Status and status inconsistency, like general and bipolar factors, are concepts which sink or swim together. The sociologist may decide to investigate them jointly . . . or to ignore them both . . . [.] but he cannot identify effects of inconsistency without first looking at effects of status.

One of House's criticisms is that this is not what Lenski, in particular, meant by status inconsistency. My contention is that, although Lenski's attitude to his original formulation was equivocal, there is no doubt that the basic idea can be taken in my sense, for he defined his "quantitative measure of status crystallization" as 100 minus "the square root of the sum of the squared deviations from the mean of the four hierarchy scores of the individual" (Lenski, 1954:407). That mean represents a single measure of general status. The first few paragraphs of Lenski's paper set out quite clearly what he was trying to do. He begins by pointing out that social philosophers from Aristotle to Marx have postulated a single, unidimensional, status hierarchy. He then goes on to

say that status has in fact several facets, and this naturally leads him, by way of a reference to Benoit-Smullyan (1944), to entertain the idea that "certain units may be consistently high or consistently low, while others may combine high standing with respect to certain status variables with low standing with respect to others" (p. 405). This is a clear reference to what I call status discrepancy, and it contains no suggestion that the status variables are interacting in the statistical sense. This interpretation of Lenski's position is supported by Sampson (1963) in a paper to which I referred in the course of my earlier exegesis.

What the proponents of the additive model did was to go one step further. Lenski had added the distinct aspects of status to the overall general factor, now they dispensed with the general factor altogether and analysed only the specific status variables. In so doing they deprived themselves of the possibility of measuring departure from status crystallization as squared deviation from mean status; that is, they threw out the idea of inconsistency as it had been originally conceived. It is true that Lenski fell in with this way of speaking, and I certainly do not deny him "the right to clarify his theory as he did in 1964," but equally I claim the right to consider the theory in its pre-1964 state. Indeed, I suspect that Lenski was not happy with the new formulation, sensing that his original conception was being mangled in some way.

In its original form inconsistency between aspects of status was a phenomenon of *verstehen*. As Benoit-Smullyan (1944:160) said

the economic progress of the bourgeoisie [in the ancien régime] would not in itself have produced the ensuing tension, were it not for the fact that there existed strong psychological and social forces working towards status equilibrium which were continually blocked or repressed by institutional barriers. Unless men expected power and prestige to accord with wealth, there would be no frustration involved in a disequibrated status structure.

This point, in fact, disposes of House's main argument. My claim was that a status inconsistency effect is, at its simplest, a status difference effect; that is, a situation such that the net regression coefficient for one variable is greater than the net regression coefficient for the other. Speaking more precisely, what I demanded was that the empirically observed regression coefficients should deviate significantly from a set of *a priori* weights which have been applied to the independent variables in order to define a measure of overall status. House interprets this stipulation as a claim that all status difference effects are status inconsis-

tency effects. But this is to turn a theoretical derivation into an induction; status inconsistency cannot plausibly be presented as the only possible reason why status difference effects occur. It is incumbent on the theorist of status inconsistency to prescribe the circumstances in which a discrepancy between status dimensions is (in some sense) perceived or experienced as a state of normative disarray, that is, as inconsistency. Status inconsistency is a discrepancy between positions on two (or more) axes on which people are expected, or would prefer, to be equal (Anderson and Zelditch, 1964).

One of the reasons why the additive model caught on among empirical sociologists is that, by defining inconsistency as interaction rather than discrepancy (that is, as a difference between differences, rather than as a difference *simpliciter*) it enabled them to test for the presence of inconsistency effects without delving into people's minds. The underlying conceptual shift went like this: inconsistency is a discrepancy which creates social discomfort; we have no data on discomfort so we are left with discrepancy; but this is too simple and also too common to be theoretically interesting; so we shall try defining inconsistency as interaction; and when we do that we find that inconsistency (and mobility) have virtually no social consequences or concomitants; which means that people manage to accommodate themselves pretty well to whatever combination of conditions they find themselves in.

The shift did not pass unnoticed. Meyer and Hammond (1971), for example, devoted a paper to examining the congruence rules which define status inconsistency. These rules are part of the normative baggage of the general public, and the sociologist should be wary of efforts to identify inconsistency with a concept (statistical interaction) which he himself did not entirely comprehend until he entered graduate school. The diamond model gives a speaking part to the social actor and invites him to say what he thinks inconsistency is. In doing this it is bringing theory back into the play, without expelling quantification.¹ That is

¹ In footnote 19 of my paper I took issue with the reductionist psychologism of writers who wish to substitute the study of *verstehen* for the study of objectively defined inconsistency. My point is that *both* are required; we must demonstrate that a discrepancy between status axes is significantly associated with the dependent variable, and we must also provide a reason for supposing that the observed discrepancy is perceived or experienced as inconsistency. The problem is similar to that faced by reference group theory when it is applied to the study of

what I meant by saying that the model is analytic rather than merely descriptive.

House quite rightly says that my quarrel with the additive model is "not really methodological or empirical but rather theoretical and conceptual." But he is wrong in alleging that "he prefers the composite status dimension because it is derived (e.g., by statistical procedures such as factor analyses) rather than directly observed." On the contrary, I list four possible ways of arriving at a measure of general status, two of which involve measuring it directly (Hope, 1975:327f).

Lenski entered the stratification field at a time when Warner and Centers had developed measures of overall status. Taking a hint from Weber via Parsons, and also from Benoit-Smullyan, he thought that it might be interesting to examine status differences and status discrepancies, as well as status sums or averages. So he instituted a geometrical model which he proceeded to collapse along the vertical dimension of general status, thus locating inconsistency in the subspace which is orthogonal to the vertical axis. The additive modelers, in effect, substituted an entirely different model. The space of both models is the same, but the primary axes of Lenski's space are fixed by theory and need not correspond to any measured variable, whereas the only axes admitted to the additive model are the measured variables. Confusion arose because the additive model was presented as a test of Lenski's hypothesis, whereas it is an alternative conceptual scheme in which that hypothesis cannot be stated. Why do I think that we should revert to a model which includes both overall status and status inconsistency?

One reason is that sociologists, and people in general, seem to believe that inconsistency can be an important motivating or debilitating force. The negative findings in the literature are beside the point because they are mostly based on applications of the additive model, and that model partials out important segments of the inconsistency variance.

The other main reason is that strong theories are to be preferred to weak ones. In applying status inconsistency theory, we have a relatively clear idea of just what it is we are holding constant and what it is we are testing. This is not true of the additive model, which is a model

for which no theory has yet been found (though I did in fact suggest a possible theoretical rehabilitation of the model towards the end of my paper). It is true that strong theory demands a high level of measurement. But quantitative analysts have not been slow to compute correlation coefficients for status axes, and the product-moment coefficient presupposes equal-interval measurement. The only further step required by inconsistency theory is identification of homologous points across the axes.²

The weakness of the linear additive model is that it runs together two stages of empirical investigation; namely, the specification of what shall count as inconsistency and the test of whether or not it occurs. The nonlinear additive model is weaker still: the rows of the table of means could be placed in any order, and the columns in any other order, without affecting the conclusions of the analysis in any way. Indeed the model is perfectly appropriate to a table whose rows represent one principle of classification while the columns represent another.

The strength of inconsistency theory, properly understood, is its postulation of a general factor of status as a foil to various aspects of status. There is a certain irony in my position as a defender of such a general axis in the field of status inconsistency, because in the field of occupational status I appear as a defender of specific aspects of status. Investigations of the prestige of occupations almost invariably confine themselves to measuring one dimension of occupational standing, namely, prestige, whereas it is possible to demonstrate that occupations are ranked in markedly different ways on distinct dimensions, of which prestige is the general factor (Hope, 1978). In both areas of study my point is really the same; viz., that it is theoretically desirable, and empirically productive, to keep both the general and the specific factors in play.

The debate is curiously reminiscent of another dispute which occurred fifty years ago among psychologists. On the one hand were those psychologists, mostly British, who argued for the existence of a general factor of intelligence; on the other hand were those,

relative deprivation. It also resembles the problem of what counts as a reward in reinforcement theory.

In the printing of footnote 19 most of the first sentence below the line on p. 338 was omitted. I take this opportunity of giving the sentence in full. "Doreian and Stockman are wrong in supposing that a basic assumption of status inconsistency theory is that every variable has a fixed number of ranks."

² House is quite wrong in suggesting (fn. 3) that I fail to observe that axes must be in the same metric if they are to be added or subtracted. The problem of scaling is alluded to at several points in my paper; indeed Lenski's call for common scales is quoted in the third paragraph. Scaling is taken so seriously that four possible implementations of the diamond model are presented (p. 334f) so that the analyst may choose a version which does not place more weight on his metrics than they can bear.

many of them American, who held that it was desirable to work with a set of oblique factors. (Oddly enough, British social scientists have almost given up work on intelligence, while American sociologists are now engaged in theoretical and empirical work on the general factor of IQ.) The debate was a nonissue. Insofar as measures of cognitive competence are all positively correlated there is every reason to study their general factor. Insofar as they are not perfectly correlated there is equally good reason to study the discrepancies among them. The latter is more difficult than the former because difference measures are less reliable than sums or averages (a point which I made in my paper). The fact that these issues are stated in the jargon of factor analysis should not mislead us into supposing that we are condemned to employ that technique. It is much to be preferred that we should set out to find direct, unbiased measures of our theoretical constructs. The very idea that inconsistency is a normative phenomenon carries with it the implication that people should be able to tell us what relations between status axes count as status equilibrium. In other words, if there is anything in status inconsistency theory, we ought to be able to elicit a measure of general status from the man in the street, just as we can for the prestige of occupations.

The square-additive model, which does not involve a prior identification of general status, has been employed in several score research reports over the last ten or twelve years. It has been uniformly successful in achieving a good fit to data on many variables; even where deviations from the model are significant they are negligible in degree. If we were to accept these results at face value we should have to say that status inconsistency (and also social mobility, for the model has applied to both phenomena without distinction) is a sociologically inert condition, having no consequences or concomitants, either good or bad. But my contention is that the additive model is an alternative to, not a test of, status inconsistency theory, and hence that the substantive issue is still open.

The diamond model was introduced in order to do justice to two different research traditions by unifying them in a single approach which incorporates the strong points of both. One of those traditions, namely, the formulation of "congruency rules," or statements of what shall count as status consistency and status inconsistency, has latterly been subsumed by the other, namely, the additive model. The former is strong on meaning and weak on testability, while the latter is strong on

testability and weak on meaning. The diamond model does not purport to be the only possible representation of inconsistency theory in all its possible forms. On the contrary, the general recommendations which were made for exploring the structure of a model by means of "design matrix regression analysis" (Hope, 1975:337f) were an attempt to introduce a more flexible approach to the specification of theory. What is important about the model is that, by setting in train the break-up of a research consensus, it should induce sociologists to clarify and elaborate the theoretical basis of their investigations as they mull over the substantive significance of nonnegligible deviations from the model. Their new conceptions will inevitably impel them to construct new models.

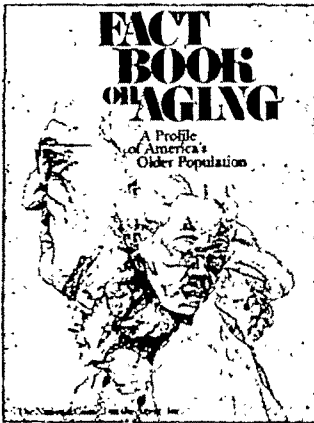
In his analysis of what he calls "problemshifts," Lakatos (1970) distinguishes between progressive and degenerating problemshifts. When it was first introduced the additive model was undeniably progressive. Partial and incomplete it may have been, but it was a determinate model, whose properties were formally specified. Being a proper model, it could be set to work, and taken apart when it broke down. Unfortunately it hardly ever broke down, so it was never taken apart. As a consequence research on status inconsistency stuck in a degenerating phase. In my paper I showed that the empirical omniscience of the additive model is a consequence of the fact that the model has the most important aspects of the diamond model up its sleeve, and I argued that it is time it put its cards on the table.

Keith Hope
Nuffield College, Oxford

REFERENCES

- Anderson, B. and M. Zelditch
1964 "Rank equilibration and political behavior." *Archives Européennes de Sociologie* 5:112-25.
- Benoit-Smullyan, E.
1944 "Status, status types, and status interrelations." *American Sociological Review* 9:151-61.
- Hope, K.
1975 "Models of status inconsistency and social mobility effects." *American Sociological Review* 40:322-43.
1978 "What are people doing when they grade occupations?" Mimeo. Nuffield College, Oxford.
- House, J. S.
1978 "Facets and flaws of Hope's diamond model (comment on Hope, ASR June, 1975)." *American Sociological Review* 43:439-42.

- Kuhn, T. S.
1970 *The Structure of Scientific Revolutions*. 2nd ed. Chicago: University of Chicago Press.
- Lakatos, I.
1970 "Falsification and the methodology of scientific research programmes." Pp. 91-196 in I. Lakatos and A. Musgrave (eds.), *Criticism and the Growth of Knowledge*. Cambridge, Eng.: Cambridge University Press.
- Lenski, G.
1954 "Status crystallization: a non-vertical dimension of social status." *American Sociological Review* 19:405-13.
- Meyer, J. W. and P. E. Hammond
1971 "Forms of status inconsistency." *Social Forces* 50:91-101.
- Sampson, E. E.
1963 "Status congruence and cognitive consistency." *Sociometry* 26:146-62.



Fact Book on Aging

A Profile of America's Older Population

Published by the National Council on the Aging. 270 pp., 8½ x 11, paperback. \$8.50, NCOA Members; \$10, Non-Members.

Invaluable reference tool.

First complete sourcebook of information on U.S. older population. Consolidates extensive current data. Categorizes and summarizes present knowledge in eight areas of aging: demography, income, employment, physical and mental health, housing, transportation and criminal victimization.

Tables, charts and bibliography accompany the narrative of each chapter.

Provides total profile, filling out easily accessible facts with data not readily available.

Sources include computer tapes, national social science studies, books and articles by leading gerontologists, federal government publications and the landmark NCOA/Louis Harris & Associates survey, *The Myth and Reality of Aging in America*.

Order from: **Publications Dept. S, The National Council on the Aging, Inc.**
1828 L St., NW, Washington, D.C. 20036

ITEMS (Continued)

aspects of environmental issues. ROBERT BAUMGARTNER is a graduate student in the Department of Rural Sociology at the University of Wisconsin, Madison. His research centers on substitutability in outdoor recreation.

■ WILLIAM D. CRANO (Expressive and Instrumental Role Complementarity in the Family) is Professor in the Department of Psychology at Michigan State University. He is studying the effects of socioeconomic status on teacher expectations and educability, methodological approaches allowing for a refinement of currently popular causal inference techniques, and cognitive social psychology. He has authored (with Marilyn B. Brewer) "Fundamentos de la investigación en psicología social," *Editorial El Manual Moderno*, 1977. JOEL ARONOFF is Professor in the Department of Psychology at Michigan State University. His research interests include personality influences on group structure and functioning.

■ JANET LEVER (Sex Differences in the Complexity of Children's Play) is Assistant Professor in the Department of Sociology at Northwestern University. She is working on a comparative study of the role of sport in modern society with Brazil and the United States as the chief cases.

■ MARGARET MOONEY MARINI (Transition to Adulthood) is Research Scientist at Batelle Human Affairs Research Centers. Currently she is investigating sex differences in educational, occupational, and income attainment and the implications of these attainments for social standing and well-being.

■ JOHN F. MYLES (Sick Role Identification among the Elderly) is Assistant Professor in the Department of Sociology and Anthropology at Carleton University. His research focuses on the political economy of income maintenance programs for the aged and the role of these programs in reproducing and altering the structure of intercohort and intra-cohort inequality.

■ WILLIAM A. RUSHING (Status Resources, Societal Reactions, and Mental Hospital Admission) is Professor and Chairperson of the Department of Sociology at Vanderbilt University. His research interests include the social and legal aspects of mental hospitalization and release, and the consequences of illness on individual behavior and family functioning. In 1975 he authored *Community, Physicians, and Inequality: A Sociological Analysis of the Maldistribution of Physicians* (D. C. Heath). He is also coeditor (with Mayer N. Zald) of *Organizations and Beyond: The Collected Essays of James D. Thompson* (D. C. Heath, 1976).

■ PATRICK M. HORAN (Is Status Attainment Research Atheoretical?) is Associate Professor of Sociology at the University of Georgia. He is pursuing the stratification research implications of dual economy theory, with special emphasis on the analysis of discrimination and intergenerational mobility.

■ LYNN SMITH-LOVIN (Labor Force Participation, Fertility Behavior, and Sex Role Attitudes) is a Ph.D. Candidate in the Department of Sociology at the University of North Carolina, Chapel Hill. Her current research (with Ann R. Tickamyer) is a cross-national study of the relationship between female labor force participation and fertility levels. She also is studying voluntary childlessness in the United States and the effect of behavioral settings on affective reactions to social events. ANN R. TICKAMYER is a Ph.D. Candidate in the Department of Sociology at the University of North Carolina, Chapel Hill. Besides the work she is doing with Smith-Lovin, she is studying property stratification and decision making among elites.

■ GLENN FIREBAUGH (Individual-Level Relationships from Aggregate Data) is Assistant Professor in the Department of Sociology and Anthropology at Vanderbilt University. Currently he is studying the causes of agricultural depopulation in Western Europe and North America.

■ DENNIS P. HOGAN (Order of Events in the Life Course) is Assistant Director and Research Associate in the Community and Family Study Center at the University of Chicago. He is engaged in fertility research in Thailand and Egypt. He continues to do research on racial and ethnic differentials.

■ DOUGLAS S. MASSEY (Comment on Cortese, Falk and Cohen, ASR August, 1976) is a Ph.D. Candidate in sociology at Princeton University. He is examining a method of measuring regional population density using contiguity methods. He also is studying residential segregation among nonlabor force populations.

■ W. DAVID PIERCE (Comment on Burgess and Nielsen, ASR June, 1974) is Assistant Professor of Sociology at the University of Alberta. His current research interests include reward schedules in social exchange; interpersonal attraction as a function of pattern and proportion of agreement; and performance predictions and withdrawal of evaluations and social comparisons. W. ANDREW HARRELL is Associate Professor of Sociology at the University of Alberta. He does experimental work in social exchange theory and field experimental research in bystander intervention following the occurrence of a theft.

■ MARK STUMPP (Comment on Chase-Dunn, ASR December, 1975) is a Ph.D. Candidate and Teaching Associate in the Department of Economics at Brown University. He is investigating the determinants of corporate financial structure and their effects on firm behavior. ROBERT M. MARSH is Professor of Sociology at Brown University. He is interested in modernization and development. His book (coauthored with Hiroshi Mannari), *Modernization and the Japanese Factory*, was published by Princeton University Press in 1976. DEBORAH A. LAKE received her A.B. from Brown University.



MOBILIZING WOMEN FOR WAR
German and American Propaganda, 1939-1945

LEILA J. RUPP

To discover how war can affect the status of women in industrial countries, Leila Rupp examines mobilization propaganda directed at women in Nazi Germany and the United States. Her book explores the relationship between ideology and policy, challenging the idea that wars improve the status of women by bringing them into new areas of activity. Illus. \$12.50

HISTORICAL STUDIES OF
CHANGING FERTILITY

Edited by CHARLES TILLY

These papers examine the historical experience of particular populations in Western Europe and North America in a search for the processes that change fertility patterns. "... A splendid survey of demographic history and its importance to the understanding of the historical processes that have characterized western society since before the Industrial Revolution."

— Marc Nerlove, *Northwestern University*
Quantitative Studies in History

Cloth, \$27.50 • Limited Paperback
 Edition, \$10.95

THE STATUS OF WOMEN IN
PREINDUSTRIAL SOCIETIES

MARTIN KING WHYTE

Whyte examines how the status of women in different cultures compares with that of men and varies within religious, political, economic, domestic, and sexual contexts, by drawing on a cross-cultural sample of 93 preindustrial societies throughout the world. His analysis describes women's roles in historical perspective, offering a much-needed foundation for feminist scholarship as well as provocative thoughts about the future. \$15.00

THE STATE AND SOCIETY
Peru in Comparative Perspective

ALFRED STEPAN

Although the state's role in society has clearly expanded since the 1930s, its independent effect on social structure and change has been given little weight in modern political theories. To bring theory more into line with reality, Stepan proposes a new model of state autonomy which he shows to be particularly well suited for understanding political developments in the Iberian countries and their former Latin-American colonies.

Cloth, \$18.50 • Paper, \$4.95

Princeton University Press

Princeton, New Jersey 08540

NEW FROM
COLUMBIA

Coming this Fall . . .

WHAT IS SOCIOLOGY?

Norbert Elias

This unique book provides a concise introduction to the thought of Norbert Elias, long recognized as a major sociologist in France and Germany and now attracting deserved attention in the U.S. In *What Is Sociology?* Elias refines his assertions about the making of modern man and modern society which were first argued in his massive work on the civilizing process. \$17.50

HERMENEUTICS AND SOCIAL SCIENCE

Zygmunt Bauman

In this first English-language introduction to the German philosophical tradition of hermeneutics, Bauman skillfully illuminates the subjective factors necessary to a real understanding of social phenomena. His authoritative analysis covers such seminal figures as Weber, Mannheim, Marx, Heidegger, Schutz, and Parsons. \$15.00

THE TASK-CENTERED SYSTEM

William J. Reid

In this unparalleled introduction to social treatment the famous model for helping people, constructed in Reid and Epstein's *Task-Centered Casework*, is shown to be matured, developed, and tested in a wide range of practice contexts. For practitioners and students at any level. \$12.50

To order send check or money order to Dept. JN. Individuals must enclose payment. Institutions may request billing.



COLUMBIA UNIVERSITY PRESS

Address for orders:
136 South Broadway, Irvington, New York 10533

+ + +

MANUSCRIPTS FOR THE ASA ROSE SOCIOLOGY SERIES

Manuscripts (100 to 300 typed pages) are solicited for publication in the *ASA Arnold and Caroline Rose Monograph Series*. The Series welcomes a variety of types of sociological work—qualitative or quantitative empirical studies, and theoretical or methodological treatises. An author should submit three copies of a manuscript for consideration to the Series Editor, Professor Robin M. Williams, Jr., Department of Sociology, Cornell University, Ithaca, New York 14853.

THE HAWTHORNE EXPERIMENTS: FIRST STATISTICAL INTERPRETATION*

RICHARD HERBERT FRANKE
Worcester Polytechnic Institute

JAMES D. KAUL
Environmental Protection Agency

American Sociological Review 1978, Vol. 43 (October):623-643

A guide is provided to the proceedings of the Hawthorne experiments, and experimental data are now made readily available. Data from the main experiment (that in the first relay assembly test room at Western Electric) are interpreted statistically for the first time. Quantitative analysis of this quasi experiment is accomplished by time-series multiple regression using nearly five years of data. This analysis demonstrates that experimental variables account for some 90% of the variance in quantity and quality of output, both for the group and for individual workers. Imposition of managerial discipline, economic adversity, and quality of raw materials provide most explanation, obviating the need to draw upon less clearly definable human relations mechanisms. For decades the Hawthorne studies have provided a rationale for humane approaches in the organization of work by suggesting that considerate or participative treatment of workers led to better economic performance. The present analysis suggests, to the contrary, that humanitarian procedures must provide their own justification.

The massive Hawthorne experiments of some 50 years ago serve as the paradigmatic foundation of the social science of work.¹ The insights gleaned from these

experiments provide a basis for most current studies in human relations as well as for subareas such as participation, organizational development, leadership, motivation, and even organizational design. But aside from visual inspection and anecdotal comment,² the complex of data obtained during the eight years of the Hawthorne experiments has never been subjected to thorough-going scientific analysis. Indeed, as was pointed out in this journal by Carey (1967), the data necessary for statistical analysis are not available in the scientific literature. It is the purpose of this report to make the Hawthorne data accessible, to interpret systematically the most important of these, and to draw from the results thus obtained some conclusions regarding the use of social science in industry.

Since interpretation and criticism of the Hawthorne studies to date have been little

* Address all communications to: Richard H. Franke; Department of Management; Worcester Polytechnic Institute, Worcester, MA 01609.

We are indebted to the Western Electric Corporation, to the Library and the School of Business Administration of the University of Wisconsin, Milwaukee, and to the Worcester Polytechnic Institute for facilitating and supporting this study. For their help in obtaining the Hawthorne documents, we thank particularly E. L. Byrom and Florence Hybl of Western Electric; William Roselle, Stanley Mallach, and Steven Smith of UWM Library; and Ruth Zubrensky and George Carian. We also thank two anonymous reviewers and the following persons for their comments on earlier drafts: Bernard Bass, Robert Dubin, Elke Franke, Arthur Gerstenfeld, Donald Gerwin, Chadwick Haberstroh, Alex Inkeles, Wilbert Moore, Stephanie Pluskota, James Price, and Winston Ring. An earlier report of these findings was presented at the annual meeting of the American Association for the Advancement of Science, Washington, 1978.

¹ For a selection of studies testifying to the importance of the Hawthorne studies in the development of applied social science, cf. Homans (1941; 1950), Friedman (1946), Miller and Form (1951), Viteles (1953), Blum and Naylor (1968), Sanford (1973), Cass and Zimmer (1975), and Locke (1976).

² Such interpretation has been taken to considerable lengths, as in the work of Roethlisberger and Dickson (1939), Homans (1950), and Whitehead (1938). The latter author did also employ statistical procedures, but without application to the major dependent variables in the various experiments.

more than opinion, most of this introduction will be a simple description of the Hawthorne experiments over 1924 to 1933, with brief note of the conclusions and impact of these studies.³ Systematic review of the secondary literature is presented following the analytical section, so that these evaluations may be judged in light of the results of quantitative analysis.

The Hawthorne studies began in 1924 at the Hawthorne plant of the Western Electric Company in Chicago with an inquiry by the National Academy of Sciences and Western Electric into relationships between illumination levels and worker production rates. Inexplicably worker output and job satisfaction generally increased regardless of increase or decrease in illumination. Their curiosities piqued, Western Electric management and social scientists from the Harvard School of Business Administration initiated experiments to examine effects of social as well as physical factors upon work efficiency. A chronology of the experiments is presented in Figure 1. The exploratory illumination experiments (1924-27) were followed by the main Hawthorne experiment, in the first relay assembly test room (1927-33), and by four derivative experiments (1928-32). The first four experimental programs were reactive; that is, conditions were manipulated by the experimenters, who then noted changes in work satisfaction and performance. The final two experiments did not include advertent manipulation of independent variables. However, the presence of interviewers and observers was itself a change in the conditions of work.

A flowchart and description of events is presented in Figure 2.⁴ From sole attention to environmental conditions of work in (1) the illumination experiments, the studies expanded in (2) the first relay experiment to scrutinize effects of work en-

vironment, physical requirements, management, and social relations upon output. All issues dealt with subsequently were initiated, at least broadly, in the first relay experiment. The derivative studies were: (3) the second relay experiment, which tested and discounted effects of small group incentive payment; (4) the mica splitting experiment, which tested and discounted effects of rest pauses upon performance; (5) the interviewing program, which indicated that relations with management and with peers were important to worker satisfaction, and that informal group organization could be used by workers to regulate and reduce the pace of their work; (6) the bank wiring observation, which confirmed the latter conclusion regarding output restriction, and thus underlined the importance of social relations among workers. Counseling, supervisory training, and other nonexperimental programs also were undertaken by Western Electric to make use of the conclusions from the six experiments. In experiments (2), (3), (4), and (6), research attention focused on small group activities. Three separate groups of five female workers each were involved in the first relay, second relay, and mica splitting experiments, while 14 male workers participated in the bank wiring study. The (1) illumination and (5) interviewing studies, on the other hand, involved whole departments of workers.

The researchers concluded from both the primary and the derivative experiments that measured experimental variables had little effect, but that the unmeasured quality of human relations of workers to management and peer group was responsible for most output improvements observed in the first four experiments. This rather unspecific conclusion, providing a foundation for modern humanitarian and human relations approaches to work, led other researchers to focus upon worker satisfaction, as in the Ohio State supervision studies and their descendants (cf. Fleishman et al., 1955; Miner, 1965), to studies of authoritarianism (cf. Sales, 1966; Vroom, 1960), informal organization (Whyte, 1955), leadership (Bass, 1960; Stogdill, 1974), participative management (cf. Likert, 1967; Marrow, 1975; Pusić,

³ The primary sources describing and interpreting the Hawthorne experiments are the monumental *Management and the Worker* of Roethlisberger and Dickson (1939), two volumes of graphs and description by Whitehead (1938), and description and social application by Elton Mayo (1933; 1945; 1947).

⁴ The description and conclusion are extracted from Roethlisberger and Dickson (1939).

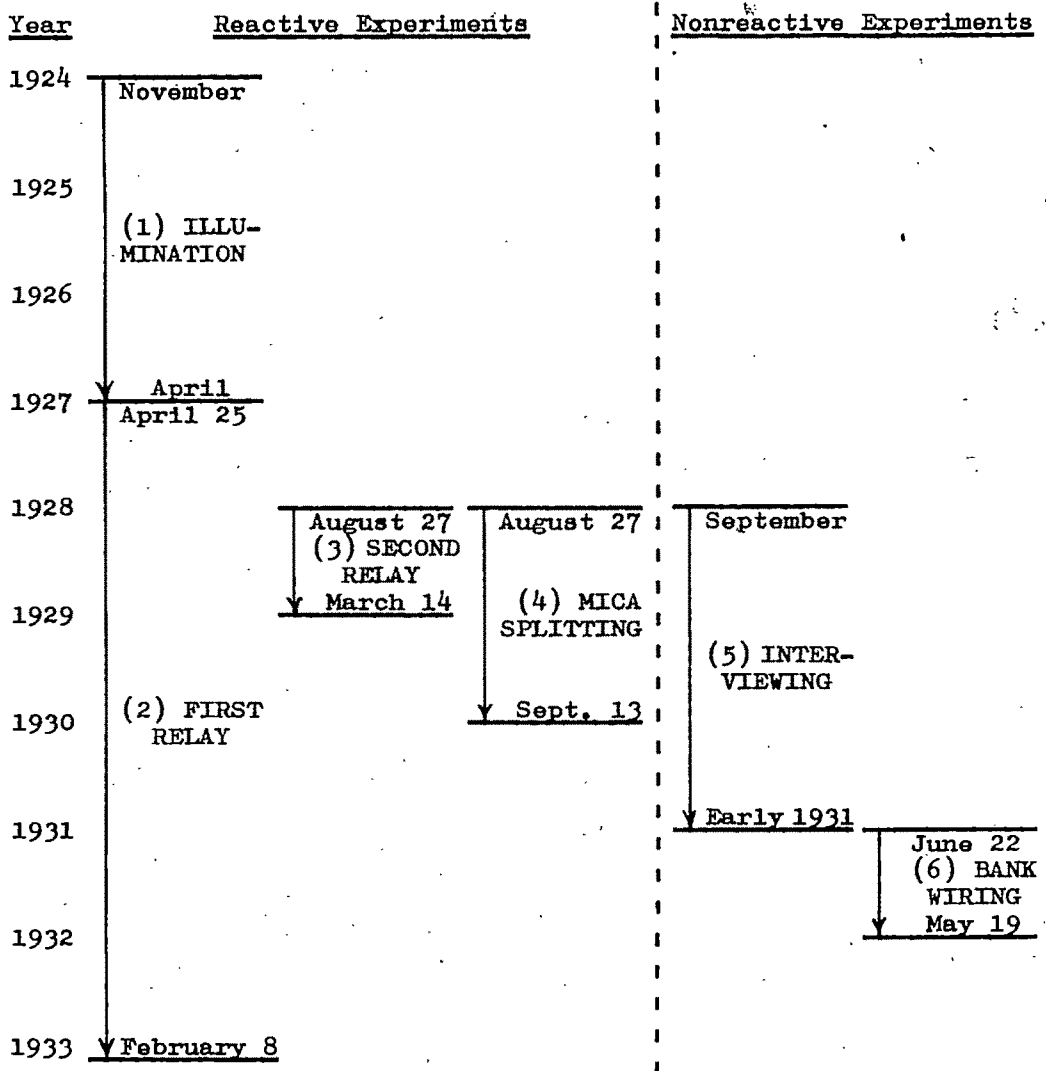


Figure 1. Chronology of the Hawthorne Experiments

1973), and to the use of sensitivity training and related techniques in organizational development (cf. Beer, 1976; Bradford et al., 1964; Golembiewski and Blumberg, 1970). It should be stated here that the initial concern of the Hawthorne experiments was with output. Concentration upon worker satisfaction in subsequent studies is sometimes justified by assuming it to be an intervening factor in job performance. This is an assumption often made in practice if not in word (as by Kahn, 1975, and Price, 1968), in the face of contradicting evidence (cf. Locke, 1976; Vroom, 1964). Attempts to demon-

strate empirically a linkage between human relations and work performance have not received primary attention since the time of the seminal Hawthorne experiments. To cast further doubt upon the human relations conclusion, we should again note that there have never been meaningful statistical analyses of the data from the Hawthorne experiments. The absence of statistical analysis may have resulted from the nature of the experiments: (a) there were no control groups other than the experimental groups themselves prior to manipulation; (b) since these were field experiments over extended periods

<u>Flowchart</u>	<u>Experimental Variables</u> <u>Manipulated or Observed</u>				<u>Description and</u> <u>Conclusion</u>
	<u>Physi- cal Work Envi- ron- ment</u>	<u>Physi- cal Work Re- quire- ments</u>	<u>Manage- ment and Super- vision</u>	<u>Social Rela- tions of Workers</u>	
(1) ILLU- MINATION	X				Three exploratory studies that suggested human factors rather than physical working conditions determined worker satisfaction and performance.
(2) FIRST RELAY	X	X	X	X	The major Hawthorne experiment, testing effects on performance of rest pauses, shorter work periods, and increased worker autonomy and of small group incentive pay. This study concluded that benefits to worker performance resulted from improved human relations, and to a lesser extent from rest pauses.
(3) SECOND RELAY			X		Derivative experiment suggesting only moderate effects of small group incentive pay upon worker performance.
(4) MICA SPLITTING		X			Derivative experiment suggesting only moderate effects of rest and shorter work periods upon worker performance.
(5) INTER- VIEWING			X	X	Derivative survey reinforcing prior conclusions regarding importance of social interactions (worker-worker and worker-supervisor) in the satisfaction of workers. First indications, during intensive interviews early in 1931, of problems resulting from employee interrelations, especially in restriction of output.
(6) BANK WIRING				X	Derivative observations noting the effectiveness of social interactions in a large group of workers in standardizing the pace of work (restricting output during period of economic depression).

Figure 2. Flow, Content, and Conclusions of the Hawthorne Studies

of time, it would have been difficult to eliminate all extraneous variables from the experiment.⁵

Fortunately, quasi-experimental approaches and time-series analytical procedures have been developed since the original researchers' interpretations of the Hawthorne experiments (Mayo, 1933; Roethlisberger and Dickson, 1939; Whitehead, 1938). These methods allow the use of periodic data, with testing for effects of measured variables and adjustment for changes over time in unmeasured historical factors. Rough schematic approaches have been suggested for analysis in education and psychology by Campbell and Stanley (1963) and Cook and Campbell (1976), while quantitative regression procedures have been developed in the field of econometrics (cf. the analytical procedures described below).

One investigation at Hawthorne, the first relay experiment, included a variety of dependent and independent variables that could be expressed quantitatively over 23 experimental periods,⁶ allowing convenient use of time-series regression. Measures of quantity and quality of output could be obtained for the group of five workers and for each individual, and measures of independent variables also could be obtained. Rest pauses, hours of work per day, and days of work per week were intentionally manipulated, and a small group incentive system was introduced. A number of inadvertent categoric changes which occurred over the five years of the experiment also could be identified. Most interesting of these changes were the replacement by management of two of the workers after period 7, because of their unsatisfactory attitudes in response to requests for greater diligence and more output, and the onslaught of the great depres-

sion early in period 15. The analyses that follow will use the available evidence to test directly for the sources of differences in worker performance over time, to determine whether the substantial performance variances obtained in the experiment can be explained quantitatively.

METHOD

Data

The original documents from the Hawthorne experiments were reviewed and then borrowed during visits to the Hawthorne plant between November 1976 and May 1977. Copies of the documents are now on microfilm in the libraries of the University of Wisconsin, Milwaukee, and the Worcester Polytechnic Institute. The data of the first relay experiment are summarized in Appendix 1 (group) and Appendix 2 (individuals). Quantity of output is recorded as net hourly and net weekly rate per worker, while quality of output is recorded as repair time required per day. Hours worked per day and per week and the number of weeks per experimental period describe the basic work schedules, with time taken for scheduled and voluntary rest pauses reducing actual work time. Categoric changes in working conditions are expressed as dummy variables of zero to one for managerial discipline (the replacement by management of two of the five workers, with one of the replacements assuming the role of straw boss), and for the occurrence of the economic depression, the supply of defective raw materials for two periods, the temporary voluntary replacement of one worker, and for the change from a large group to a small group incentive system of pay after the first two experimental periods. A list of these variables and their dimensions is provided in the first results table presented below. Data were available over the 23 periods for all variables, except for repair time (18 periods) and voluntary rest time (21 periods), each unmeasured in several early and late periods.

Analytical Procedures

The data of the first relay experiment are suitable for rigorous analysis using

⁵ The experimental and analytical approaches of the Hawthorne experimenters have been criticized in these and other lights by numerous authors, including Carey (1967), Cook and Campbell (1976), and Farris (1969), and by those reviewed by Landsberger (1958).

⁶ There was also a 24th period, from March 1, 1932, to February 8, 1933, which is not considered in the present analyses. During this final period, all five operators were laid off and replaced by more senior workers who were inexperienced in the assembly of relays.

time-series econometric techniques. With these, even influences of inadvertent experimental changes can be examined specifically.⁷ In addition, influence of other (generalized) historical factors as well as the passage of time can be measured as *serial correlation* (Durbin and Watson, 1950; 1951), and any effects upon the statistical independence of sequential sets of data can be removed (using the approach of Theil and Nagar, 1961, as described by Johnston, 1963, and Elliott, 1973).

Analyses are directed toward explanation of differences in output over time for the group, as well as for individual workers. As a first step, zero-order correlations are examined for the first relay group—for the entire 23 periods, and separately for the seven periods prior to replacement of two unsatisfactory workers and for the 16 periods after this exercise of managerial discipline. In the second step, the best multiple regressions are determined for the group using all available periods for each of the three production measures, with correction for serial correlation where necessary. As a third and final step, this procedure is repeated for the production rates of each of the individual workers, first by forcing the group models upon the individual data and then by determining whether alternate models provide greater variance explanation.⁸

RESULTS

Zero-Order Relationships for the Group

Results from correlating all group data over the experimental periods are presented in Table 1, employing the maximum number of periods available for each pair of variables. The group dependent variables are net measures of output quantity per worker per hour, and per week, as well as a negative measure of

output quality which also is subsumed in the quantitative measures (repair time required per worker per day). These three performance variables were correlated over periods with each other and with the 12 independent experimental variables in work schedules, rest pauses, and categoric changes of working conditions that occurred during the five years analyzed for the first relay experiment. Results are presented separately for the entire 23 periods, for the first seven periods with the original five workers, and for the final 16 periods with the group containing three original and two replacement workers.

Because the variables are measured for sequential periods of time and their generalized historical dependence or serial correlation has not yet been measured, it would be premature to evaluate relationships in terms of statistical significance. However, there are in some cases very substantial simple relationships between group performance and experimental variables, yielding high levels of variance explanation (R^2). For example, period-by-period differences in rate of hourly output over 23 periods can be explained in large part (25% or more) by the categoric variables of managerial discipline and economic depression, with group hourly production apparently improved by management's replacement of two workers and by the depression. Also, fewer net hours per week, more scheduled rest time, and use of small group incentives appear to have improved hourly production rates.⁹ Since most intercorrelations of

⁷ For a treatment of regression equations containing dummy variables, cf. Johnston (1963).

⁸ Zero-order correlation coefficients also have been calculated for individuals over the 23, 7, and 16 periods, and group and individual regression equations have been calculated for the first 7 and for the remaining 16 periods. No additional findings of note were obtained in these calculations, which thus are not included in the present report.

⁹ The effects of incentive system and rest pauses upon output also were shown in the data of the second relay and mica splitting experiments. (The data are presented but not analyzed by Roethlisberger and Dickson, 1939: 132, 148.) The incentive effect was tested in the second relay experiment, where a group of five operators worked in one period with the existing large group incentive arrangement, in a second period with the small group of five as basis for incentive pay, and in a final period after return to the large group incentive system. The mean rates of production per worker rose from 1,634 to 1,840 and then back to 1,531 unit components per hour. With the earlier periods serving as controls for the same workers in the next periods, t-test analysis shows a significant difference only between periods 1 and 2 ($t = 2.54$, $p < .05$). That is, there was a significant 12.6% improvement in rate of production

Table 1. Experimental Correlations—Group Outputs and Repair *

	Periods 1-23			Periods 1-7			Periods 8-23		
	Hourly Output	Weekly Output	Repair Time	Hourly Output	Weekly Output	Repair Time	Hourly Output	Weekly Output	Repair Time
Dependent Variables									
(1-G) Hourly Output, units/hr.	.197			.973			-.356		
(2-G) Weekly Output, units/wk.	.225	.103		.693	.524		.187	.036	
(3-G) Repair Time, min./day									
Independent Variables									
Work Schedule									
(4) Hours per day	-.656	.378	.117	— ^d	—	—	-.549	.702	.199
(5) Days per week	-.629	.585	-.195	—	—	—	-.624	.891	-.174
(6) Net Hours per week ^b	-.758	.489	-.155	-.964	-.878	-.778	-.691	.920	-.077
(7) Weeks per period	.394	.480	.037	.349	.426	-.661	.040	.394	.028
Rest Pauses									
(8) Scheduled Rest Stops, no./day	.261	.036	.374	.761	.626	.884	.425	-.229	.324
(9) Scheduled Rest Time, min./day	.693	.161	.427	.964	.877	.788	.425	-.229	.324
(10-G) Voluntary Rest Time, min./day	-.466	-.096	-.356	-.612	-.402	-.852	-.256	.129	-.189
Categoric Changes in Working Conditions ^c									
(11) Managerial Discipline	.887	.343	.118	—	—	—	—	—	—
(12) Economic Depression	.791	-.096	.256	—	—	—	.917	-.351	.245
(13) Defective Raw Materials	.101	.275	.806	—	—	—	-.198	.226	.863
(14) Temporary Replacement of Oper. 5	.244	.594	.651	—	—	—	.064	.567	.694
(15) Small Group Incentive	.552	.272	—	.771	.840	—	—	—	—

* All data employed are per-capita averages over all periods for which data were available (see Appendix 1); these Pearson product-moment correlations are for periods common to both variables indicated (in total 18, 21, or 23), and are uncorrected for serial correlation. ^b Hours per day times days per week, minus scheduled rest time. ^c Change indicated=1, otherwise=0. ^d Dashes indicate no data or no variance in a variable.

these independent variables are smaller than their correlations with the dependent variable, prospects appear good for multivariate explanation of a large portion of variance in hourly output by known experimental variables. Separate correlations are presented in Table 1 for group hourly output with independent variables over periods 1-7 and 8-23—before and after the imposition of managerial discipline expressed by the replacement of operators 1A and 2A with operators 1 and 2. During the early periods, more rest time (leading to fewer net hours) and use of small group incentive payment seem to have been beneficial. In later periods, the economic depression and fewer net hours appear to have benefited hourly output. These shorter-term results are consistent with correlations over the entire 23 periods, where, in addition, the apparent effect of managerial discipline is shown.

Correlations for weekly output rates are also presented in Table 1, but show strong simple relationships over 23 periods only with the temporary replacement of operator 5 and with the days (and net hours) worked per week. A greater number of hours worked per week appears to have offset lower hourly rates when viewing weekly output. During periods 1-7 more rest time (reflected in slightly fewer net hours) and the introduction of small group incentives seem to have aided weekly output, while during periods 8-23 more net hours and the replacement of operator 5 seem to have been beneficial. In sum, the replacement of operator 5, more hours per week, more rest time, and use of small group incentives seem to have benefited group weekly output. Neither managerial discipline nor the economic depression shows the strong simple relationship to weekly output seen in Table 1 for hourly output. The replacement of operator 5

seems to have been important to weekly but not to hourly output, and more net hours seem to have benefited weekly output while detracting from hourly rate of output. On the zero-order surface, only rest time and small group incentive payment seem to have been useful to both rates of output.

Correlations for the third performance variable of repair time required per day (extent of poor quality output) also are presented in Table 1. Poor quality is strongly related to the use of defective raw materials and to the temporary replacement of operator 5, for periods 3-20 and for the later periods (8-20). During the early periods (3-7), lower quality is associated with more scheduled rest stops (and the attendant fewer net hours and less voluntary rest time taken) and with shorter experimental periods. Thus poorer quality output appears to result from a combination of factors which might affect quality more or less mechanically—from poor raw materials, an inexperienced worker, and from breaks in work routine by more frequent daily rest stops and more frequent changes in work schedule and conditions.

Multiple Analysis: Group Regression Equations

Each of the three dependent variables for the group has been regressed stepwise upon the 12 independent variables identified in Table 1 (data in Appendix 1). The results are presented in Table 2 as regression equations. Also presented are the multiple coefficients of determination (R^2 and cR^2 , the latter "corrected" for the number of independent variables), the Durbin-Watson coefficients of serial correlation (DW), stepwise variance explanations, and the regression equations after correction for serial correlation, if necessary.

Differences in rates of hourly output by the first relay group are explained in model 1 through managerial discipline (79%), economic depression (an additional 14%); and through scheduled rest time (4%). Most of this 97% variance explanation appears to have resulted from the imposition of managerial discipline, which

that appeared to result from the use of a small group incentive system. Similarly for the five workers in the mica splitting test, analysis indicates a significant and even more substantial (15.5%) improvement in hourly production rate, apparently resulting from the reduction of fatigue by use of rest pauses and fewer working hours ($t = 3.34$, $p < .02$). Other positive effects of performance-contingent incentives and of fatigue reduction upon output rates have been reported in various settings (cf. Bass and Barrett, 1972; Cherrington et al., 1971; Taylor, 1911).

Table 2. Stepwise Regression Equations—Group ^a

	Hourly Output	Inter- cept	Managerial Discipline	Economic Depression	Scheduled Rest Time		
Model 1:	$X_{1-0} = 50.50 + 8.64 X_{11} + 6.18 X_{12} + 0.18 X_7$ $r_p = .928 \quad r_p = .894 \quad r_p = .758$ $R^2 = 97.10\%$, $cR^2 = 96.64\%$, $DW = 1.58$ ($p < .05$)						
Stepwise variance explan.:	^b 78.72% + 14.46% + 3.91%						
Corrected for serial correlation (factor = 0.25):							
	$X_{1-0} = 38.29 + 8.26 X_{11} + 6.34 X_{12} + 0.16 X_7$ $r_p = .894 \quad r_p = .859 \quad r_p = .692$ ($p = .0000$) ($p = .0000$) ($p = .0007$) $R^2 = 94.48\%$, $cR^2 = 93.56\%$ ($p = .0000$), $DW = 1.95$ (NS)						
	Weekly Output	Inter- cept	Net Hours Per Week	Managerial Discipline	Economic Depression	Scheduled Rest Time	Small Group Incentive
Model 2:	$X_{2-0} = -716.33 + 66.52 X_8 + 393.78 X_{11} + 245.97 X_{12} + 7.67 X_9 + 105.32 X_{13}$ $r_p = .977 \quad r_p = .942 \quad r_p = .845 \quad r_p = .731 \quad r_p = .423$ ($p = .0000$) ($p = .0000$) ($p = .0000$) ($p = .0004$) ($p = .0713$) $R^2 = 96.57\%$, $cR^2 = 95.56\%$ ($p = .0000$), $DW = 1.98$ (NS)						
Stepwise variance explan.:	23.88% + 56.48% + 7.50% + 7.97% + 0.75%						
	Repair Time	Inter- cept	Defective Raw Mtls.	Scheduled Rest Stops	Economic Depression	Weeks per Period	
Model 3:	$X_{3-0} = 17.88 + 25.33 X_{13} + 2.58 X_8 + 7.50 X_{12} - 0.35 X_7$ $r_p = .951 \quad r_p = .763 \quad r_p = .770 \quad r_p = -.724$ $R^2 = 92.38\%$, $cR^2 = 90.03\%$, $DW = 2.32$ ($p < .05$)						
Stepwise variance explan.:	^b 64.98% + 14.01% + 5.00% + 8.38%						
Corrected for serial correlation (factor = -0.40):							
	$X_{3-0} = 22.36 + 25.46 X_{13} + 3.08 X_8 + 6.97 X_{12} - 0.27 X_7$ $r_p = .964 \quad r_p = .808 \quad r_p = .821 \quad r_p = -.594$ ($p = .0000$) ($p = .0005$) ($p = .0003$) ($p = .0251$) $R^2 = 94.90\%$, $cR^2 = 93.21\%$ ($p = .0000$), $DW = 2.07$ (NS)						

^a Independent variables are presented in order of appearance in stepwise multiple regression; $n = 23$ periods for output models and $n = 18$ periods for repair model.

^b Stepwise variance explanation prior to correction for serial correlation.

^c Stepwise model independent variables included net hours per week, excluding the highly correlated hours per day and days per week (cf. fn. 10).

included better performing replacement workers as well as the disciplinary example, from the beginning of period 8.

However, the time-series data used to obtain model 1 are not statistically independent, as indicated by the Durbin-Watson coefficient obtained (which is significantly distant from the neutral point of 2.00; cf. Durbin and Watson, 1951: Table 5). That is, the passage of time and unspecified historical factors appear to have influenced the regression residuals, making the use of critical values tables questionable in testing for significance of the regression equation. Correction for serial correlation using the Theil-Nagar (1961) approach is successful, as shown by the resulting nonsignificant Durbin-Watson coefficient of serial correlation. Variance explanation after correction is 94%, and

the three independent variables show substantial and highly significant partial correlations with hourly output. Slope coefficients are altered little by correction: managerial discipline apparently resulted in a production increase of eight or nine units per worker per hour, the depression in an increase of six units, and scheduled rest time which ranged from zero to 30 minutes per day resulted in an increase of up to five units per worker per hour. As shown in Appendix 1, hourly output rose from some 50 units per worker per hour up to nearly 72 units by the end of the experiment, and most of this change is accounted for through the above slope coefficients.

Regression of weekly output of the first relay group upon experimental variables in model 2 yields as explanatory variables

net hours per week,¹⁰ then managerial discipline, economic depression, scheduled rest time, and finally the introduction of small group incentive payment. Although hours per week enters first as a regression variable, it provides only 24% variance explanation. Managerial discipline is again the major explanatory variable (56%), followed by the depression and rest time (8% each) and by small group incentive (1%), for total variance explanation of 97%. Since there is no serial correlation, the equation and each partial relationship can be tested for significance. All are substantial and highly significant, except for small group incentive ($.05 < p < .10$), which did not enter the stepwise regression for hourly output of the group. Slope coefficients in the weekly equation are comparable with those for hourly output (considering that the average work week contained 42 net hours), with additional factors of hours worked and the introduction of small group incentives. As shown in Appendix 1, weekly output ranged from about 2,100 to 3,200 units per worker, averaging some 2,600 units. Changes in net hours worked, with a range from 30.33 to 48 hours per week, could have accounted for a range of 1,175 units of weekly output per worker if not offset by changes in other variables. Managerial discipline and the economic depression could account for 394 and 246 units per week, and small group incentives for 105 units per week.

The third regression model presented in Table 2 seeks explanation of differences in quality of production as measured by repair time per worker per day. Stepwise regression for the first relay group shows explanation by the use of defective raw materials (65%), more frequent scheduled rest stops (14%), the economic depression (5%), and by fewer weeks per experimental period (8%). Total variance explanation

is 92%. The acknowledged provision of defective raw materials in periods 14 and 15 appears to be the primary source of quality difficulty during the experiment, followed by increased disturbance of work routine by more rest stops and by shorter work periods containing unchanged working conditions, all apparently aggravated through stress induced by the depression. Correction of serial correlation, over the 18 periods for which repair data were available, did not substantially change slope coefficients. Interpretation of regression slopes shows the supply of defective raw materials accounting for 25 minutes of repair time per day, scheduled rest stops (zero to six) for up to 18 minutes, the depression for seven minutes, and shorter experimental periods (two to 31 weeks) for a range of about eight minutes of repair time per day.

For each group dependent variable in Table 2, there are measured experimental variables which explain well over 90% of the variance observed in production characteristics. Most of the difference in quantitative production of the group is explained by the replacement of two mediocre workers by others who from the outset demonstrated better performance. Most of the difference in quality of production is explained by difference in quality of raw materials. However, for an understanding of the meaning of these statistical results, replications are required at the level of individual workers—to measure for workers 1A plus 1 and 2A plus 2 the effects of replacement, and to measure the effects of this disciplinary example upon workers 3, 4, and 5.

*Replication: Individual Multiple Regressions*¹¹

In the remaining tables, multiple regression results are presented interpreting the individual performance data in Appendix 2. The regressions include initially only the independent variables found useful in

¹⁰ In this as in all other models presented, the variable net hours was included for potential regression equations, but the hours per day and days per week which are its constituents were excluded. However, models which included these variables also were examined. In no case did the resulting equation possess variance explanation superior to that of the corresponding model presented in this report.

¹¹ These analyses at a lower level of aggregation may be viewed as testing to guard against the ecological fallacy in making inferences based upon group data (cf. Blalock, 1961; Dogan and Rokkan, 1969; Galtung, 1967; Robinson, 1950; Thorndike, 1939).

group models (Table 2). Wherever independent stepwise multiple regressions yield results different from the group equations, these results are presented in the table footnotes. All results presented have been corrected for serial correlation (except for two instances of incomplete correction), and the variance explanations presented have been adjusted to allow for number of independent variables.

Multiple regressions for individual rates of hourly output are presented in Table 3. For workers 1, 2, 3, and 4, all three independent variables of group model 1 are significantly related to hourly output. Independent stepwise multiple regression shows no further variables to be important for workers 1, 2, and 4, but the additional variable of small group incentive payment benefiting worker 3. When model 1 is applied to worker 5, only the economic depression relates significantly to hourly output. Independent stepwise multiple regression for worker 5 does show managerial discipline as a positive factor, with the

additional and negatively related variables of defective raw materials (X_{13}) and the number of hours worked per week (X_6). The best individual equations from Table 3 provide explanation of variance in hourly output which ranges from 96% for operators 2A and 2 down to 66% for operator 4. Managerial discipline appears to have had the greatest effect upon those most directly involved in the replacement (operators 1A plus 1 and 2A plus 2), but also is significant for operators 3, 4, and 5, with smaller slope coefficients. Depression slope coefficients are about the same for operators 1, 2, 4, and 5, and lower but still significant for operator 3. Scheduled rest time is about equally important to the hourly output of operators 1, 2, 3, and 4, but unimportant to operator 5 (for whom fewer working hours seem to have been useful). Defective raw materials appear to have had an adverse effect upon the hourly output of operator 5 only.

Model 2 results for individual weekly output rates, presented in Table 4, are

Table 3. Replication of Regression Equation 1—Individual Hourly Output (X_1) upon Independent Variables from Table 2

X_1 for Operator(s)	Intercept	Managerial Discipline	Economic Depression	Scheduled Rest Time	Variance Explanation	Durbin-Watson Coefficient
1A + 1 ^a	29.67	+ 10.84 X_{11} $r_p = .840$ ($p = .0000$)	+ 7.59 X_{12} $r_p = .758$ ($p = .0001$)	+ 0.19 X_6 $r_p = .583$ ($p = .0070$)	87.15%	1.67
2A + 2A ^b	49.60	+ 13.65 X_{11} $r_p = .939$ ($p = .0000$)	+ 8.07 X_{12} $r_p = .874$ ($p = .0000$)	+ 0.18 X_6 $r_p = .641$ ($p = .0017$)	96.42%	1.67
3 ^b	51.63	+ 6.35 X_{11} $r_p = .835$ ($p = .0000$)	+ 2.74 X_{12} $r_p = .591$ ($p = .0048$)	+ 0.21 X_6 $r_p = .749$ ($p = .0001$)	91.12%	1.74
4 ^c	21.78	+ 6.55 X_{11} $r_p = .637$	+ 5.78 X_{12} $r_p = .595$	+ 0.17 X_6 $r_p = .570$	65.79%	1.33
5 ^d	20.83	+ 2.67 X_{11} $r_p = .215$ ($p = .3621$)	+ 7.52 X_{12} $r_p = .532$ ($p = .0157$)	+ 0.06 X_6 $r_p = .164$ ($p = .4885$)	28.33%	1.82

^a After correction for serial correlation using factor of 0.40.

^b Independent multiple regression yields:

$X_{1-4} = 53.49 + 5.81 X_{11} + 0.14 X_6 + 2.96 X_{12} + 3.89 X_{13}$, $cR^2 = 93.53\%$,
 $r_p = .853$ $r_p = .632$ $r_p = .688$ $r_p = .556$ DW = 1.83 (NS)
 ($p = .0000$) ($p = .0028$) ($p = .0008$) ($p = .0109$)

^c After correction for serial correlation using factor of 0.60; not fully corrected.

^d After correction for serial correlation using factor of 0.60; independent multiple regression yields after correction using factor of 0.40:

$X_{1-5} = 41.45 - 7.95 X_{11} + 4.89 X_{12} - 0.35 X_6 + 3.27 X_{13}$, $cR^2 = 85.19\%$,
 $r_p = -.820$ $r_p = .661$ $r_p = -.670$ $r_p = .509$ DW = 1.88 (NS)
 ($p = .0000$) ($p = .0021$) ($p = .0017$) ($p = .0261$)

Table 4. Replication of Regression Equation 2—Individual Weekly Output (X_2) upon Independent Variables from Table 2

X_2 for Opera- tor(s)	Inter- cept	Net Hours Per Week	Managerial Discipline	Economic Depression	Scheduled Rest Time	Small Group Incentive	Variance Explained	Durbin- Watson Coef.
1A+1 ^a	-661.18	+ 72.68 X_6	+516.73 X_{11}	+320.25 X_{12}	+ 10.15 X_9	+103.24 X_{13}	91.59%	1.99
		$r_p = .963$ ($p = .0000$)	$r_p = .884$ ($p = .0000$)	$r_p = .761$ ($p = .0002$)	$r_p = .695$ ($p = .0014$)	$r_p = .220$ ($p = .3814$)		
2A+2 ^b	-1173.33	+ 77.46 X_6	+635.97 X_{11}	+361.84 X_{12}	+ 9.98 X_9	+120.08 X_{13}	93.40%	1.97
		$r_p = .963$ ($p = .0000$)	$r_p = .942$ ($p = .0000$)	$r_p = .829$ ($p = .0000$)	$r_p = .676$ ($p = .0021$)	$r_p = .265$ ($p = .2880$)		
3	-641.62	+ 66.26 X_6	+277.74 X_{11}	+128.22 X_{12}	+ 7.65 X_9	+192.76 X_{13}	95.44%	2.00
		$r_p = .975$ ($p = .0000$)	$r_p = .891$ ($p = .0000$)	$r_p = .627$ ($p = .0053$)	$r_p = .717$ ($p = .0008$)	$r_p = .561$ ($p = .0154$)		
4 ^c	-284.29	+ 69.22 X_6	+310.10 X_{11}	+240.27 X_{12}	+ 9.10 X_9	+ 5.42 X_{13}	90.28%	1.44
		$r_p = .962$	$r_p = .688$	$r_p = .598$	$r_p = .654$	$r_p = .012$		
5 ^d	-19.70	+ 53.09 X_6	+109.48 X_{11}	+257.13 X_{12}	+ 2.57 X_9	+ 51.99 X_{13}	70.57%	1.96
		$r_p = .874$ ($p = .0000$)	$r_p = .218$ ($p = .3848$)	$r_p = .470$ ($p = .0490$)	$r_p = .161$ ($p = .5239$)	$r_p = .075$ ($p = .7686$)		

^a After correction for serial correlation using factor of 0.40; independent multiple regression yields after correction using factor of 0.40:

$$X_{2-(1A+1)} = -667.78 + 72.72 X_6 + 520.79 X_{11} + 10.36 X_9 + 320.45 X_{12}, cR^2 = 91.69\%,$$

$$r_p = .962 \quad r_p = .881 \quad r_p = .695 \quad r_p = .753 \quad DW = 1.91 \text{ (NS)}$$

$$(p = .0000) \quad (p = .0000) \quad (p = .0010) \quad (p = .0002)$$

^b After correction for serial correlation using factor of 0.10; independent multiple regression yields:

$$X_{2-(2A+2)} = -1380.15 + 78.36 X_6 + 652.06 X_{11} + 10.93 X_9 + 365.12 X_{12}, cR^2 = 94.26\%,$$

$$r_p = .961 \quad r_p = .951 \quad r_p = .751 \quad r_p = .833 \quad DW = 1.82 \text{ (NS)}$$

$$(p = .0000) \quad (p = .0000) \quad (p = .0001) \quad (p = .0000)$$

^c After correction for serial correlation using factor of 0.60; not fully corrected. Independent multiple regression yields after incomplete correction using factor of 0.60:

$$X_{2-4} = -284.06 + 69.22 X_6 + 309.84 X_{11} + 9.09 X_9 + 240.20 X_{12}, cR^2 = 90.85\%, DW = 1.44$$

$$r_p = .962 \quad r_p = .688 \quad r_p = .654 \quad r_p = .598$$

^d After correction for serial correlation using factor of 0.60; independent multiple regression yields after correction using factor of 0.40:

$$X_{2-5} = 124.38 + 48.30 X_6 - 337.95 X_{11} + 190.72 X_{12} + 174.28 X_{13}, cR^2 = 89.13\%,$$

$$r_p = .944 \quad r_p = -.813 \quad r_p = .620 \quad r_p = .586 \quad DW = 1.86 \text{ (NS)}$$

$$(p = .0000) \quad (p = .0000) \quad (p = .0046) \quad (p = .0083)$$

similar to the regression results of model 1 for individual hourly output. In addition, net hours worked per week are positively associated with output for all five workers, but the model 2 variable of small group incentive payment has a significant effect only for worker 3 (see fns. to Table 4). The best individual equations show variance explanation of weekly output which ranges from 95% for operator 3 to 89% for operator 5.

Model 3 regressions of repair time (poor quality output) upon experimental variables are presented in Table 5 for individual workers. Although the group regression equation is not fully supported by any individual results, the main variable,

the provision of defective raw materials, is significant and the most important variable for each individual. More frequent scheduled rest stops seem detrimental to production quality for three workers, while the economic depression and fewer weeks per experimental period seem to have contributed to poor quality for two workers. Independent stepwise multiple regression for individuals shows a negative influence of voluntary rest time upon work quality for operators 1A plus 1, negative influence of scheduled rest time for operator 3 (in place of number of scheduled rest stops), and a beneficial influence of managerial discipline upon work quality for operator 5. Best equations show vari-

Table 5. Replication of Regression Equation 3—Individual Repair Time (X_4) upon Independent Variables from Table 2

X_4 for Operator(s)	Intercept	Defective Raw Mtls.	Scheduled Rest Stops	Economic Depression	Weeks per Period	Variance Expln.	Durbin-Watson Coef.
1A+1 ^a	35.00 +	28.85 X_{12}	+ 2.00 X_8	+ 3.42 X_{12}	- 0.36 X_7	74.74%	2.06
		$r_p = .892$ ($p = .0000$)	$r_p = .400$ ($p = .1563$)	$r_p = .316$ ($p = .2716$)	$r_p = -.452$ ($p = .1046$)		
2A + 2A ^b	21.02 +	17.46 X_{12}	+ 2.72 X_8	+ 4.47 X_{12}	- 0.30 X_7	53.38%	2.05
		$r_p = .738$ ($p = .0017$)	$r_p = .544$ ($p = .0359$)	$r_p = .370$ ($p = .1750$)	$r_p = -.424$ ($p = .1157$)		
3 ^c	21.34 +	19.98 X_{12}	+ 1.85 X_8	+ 12.76 X_{12}	- 0.05 X_7	96.16%	2.09
		$r_p = .957$ ($p = .0000$)	$r_p = .695$ ($p = .0058$)	$r_p = .953$ ($p = .0000$)	$r_p = -.137$ ($p = .6414$)		
4 ^d	6.44 +	14.02 X_{12}	+ 1.07 X_8	+ 10.21 X_{12}	- 0.25 X_7	52.44%	1.92
		$r_p = .725$ ($p = .0034$)	$r_p = .293$ ($p = .3096$)	$r_p = .664$ ($p = .0096$)	$r_p = -.481$ ($p = .0813$)		
5 ^e	9.18 +	46.57 X_{12}	+ 5.58 X_8	+ 4.85 X_{12}	- 0.50 X_7	92.08%	1.92
		$r_p = .962$ ($p = .0000$)	$r_p = .865$ ($p = .0001$)	$r_p = .348$ ($p = .2231$)	$r_p = -.774$ ($p = .0012$)		

^a After correction for serial correlation using factor of -0.25 ; independent multiple regression yields:

$$X_{4-(1A+1)} = 37.00 + 28.36 X_{12} - 0.77 X_{10-1} - 0.33 X_7, cR^2 = 73.54\%,$$

$$r_p = .868 \quad r_p = .579 \quad r_p = -.503 \quad DW = 1.99 \text{ (NS)}$$

$$(p = .0000) \quad (p = .0187) \quad (p = .0469)$$

^b Independent multiple regression yields:

$$X_{4-(2A+2A)} = 22.41 + 15.28 X_{12} + 3.32 X_8, cR^2 = 56.20\%, DW = 2.23 \text{ (NS)}$$

$$r_p = .732 \quad r_p = .583$$

$$(p = .0013) \quad (p = .0177)$$

^c After correction for serial correlation using factor of $-.80$; independent multiple regression yields:

$$X_{4-3} = 13.90 + 20.35 X_{12} + 13.08 X_{10} - 0.28 X_7 + 0.21 X_8, cR^2 = 91.67\%,$$

$$r_p = .937 \quad r_p = .915 \quad r_p = -.683 \quad r_p = .596 \quad DW = 2.04 \text{ (NS)}$$

$$(p = .0000) \quad (p = .0000) \quad (p = .0050) \quad (p = .0191)$$

^d After correction for serial correlation using factor of 0.25 ; independent multiple regression yields:

$$X_{4-4} = 8.51 + 12.68 X_{12} + 8.14 X_{12}, cR^2 = 58.89\%, DW = 2.34 \text{ (NS)}$$

$$r_p = .689 \quad r_p = .662$$

$$(p = .0022) \quad (p = .0038)$$

^e After correction for serial correlation using factor of 0.40 ; independent multiple regression yields after correction using factor of -0.40 :

$$X_{4-5} = 34.37 + 49.09 X_{12} + 4.68 X_8 - 11.74 X_{12} - 0.48 X_7 + 8.73 X_{12}, cR^2 = 97.76\%$$

$$r_p = .990 \quad r_p = .877 \quad r_p = -.887 \quad r_p = -.772 \quad r_p = .865 \quad DW = 2.17 \text{ (NS)}$$

$$(p = .0000) \quad (p = .0001) \quad (p = .0001) \quad (p = .0020) \quad (p = .0001)$$

ance explanations for repair time ranging from 98% for operator 5 down to 56% for operators 2A plus 2.

In general, the group output models (Table 2) are the best models also for individuals (Tables 3, 4, and 5), or are within a few percentage points of amended models in variance explanation. Median variance explanations for individuals using both the group and the best models are 87% for hourly output, 92% for weekly output, and 75% for repair time—somewhat lower than variance explanations for the group of

94%, 96%, and 93%. For quantity of output, the managerial intervention in replacing two workers and the advent of the economic depression were important to all individuals as well as to the group. Both factors may be viewed as exerting certain pressures upon the workers. For quality of output, quality of raw materials was of primary importance to all individuals and to the group.

These results differ starkly from most earlier descriptions of the findings of the Hawthorne experiments. The following

section interprets the present results in juxtaposition to previous interpretations.

DISCUSSION

Multiple regression analyses over 23 periods of the first relay experiment at Hawthorne show that three variables—managerial discipline, the economic adversity of the depression, and time set aside for rest—explain most of the variance in quantity of output for the group and generally for individual workers: Two workers who exhibited undue independence from management (but, as shown in Appendix 2, did not have the lowest average production rates in the group) were replaced by two more agreeable workers. This exercise of managerial discipline seems to have been the major factor in increased rates of output for the now altered group, including increased production by the three individuals remaining. It may be speculated that improvement resulted from the positive example of the two new workers, as well as from the aversive effects of management's disposal of two of the original workers. Clear support is given to the suggestion by Carey (1967) that this intervention was a key part of the first relay experiment. As pointed out by Argyle (1953: 100), the Hawthorne researchers had provided "no quantitative evidence for the conclusion for which this experiment is famous—that the increase of output was due to a changed relation with supervision." Quantitative evaluation now does provide such evidence. However it is not "release from oppressive supervision," as suggested by Landsberger (1958:53), but its reassertion that explains higher rates of production.

Regarding the second independent variable resulting from the present analyses, the Hawthorne researchers as well as Argyle (1953) and Landsberger (1958) recognized that the economic depression beginning October 24, 1929, might have been a disturbing factor in the experiments. Yet they did not appear to suspect its positive influence on production. The increased importance of jobs and the real danger of losing them, because of the depression, may explain its positive contribution to

output quantity for the group and for all individuals.

The third dimension resulting from these analyses is worker fatigue. In his review of the Hawthorne studies and of the first several decades of criticism and application, Landsberger (1958) agreed with the conclusion of the Hawthorne researchers that reduction of fatigue did not play much of a role in the first relay experiment. This seemed indicated to them by examination of individual work rates over time and by the findings from the mica splitting experiment. However, Argyle (1953) and Carey (1967) reviewed the same evidence and came to opposite conclusions. Their interpretations are supported by the present analyses, suggesting that physical or mental fatigue reduction through rest pauses also contributed to higher output rates for the group and for four of the five workers. In the case of worker 5, fatigue reduction by working fewer hours in the week appears to have increased the rate of hourly output. Crude analysis of data from the mica splitting experiment further suggests that reduction of fatigue is beneficial for production (fn. 9).

An additional variable found related to production is the use of an incentive pay system based upon the output of the small group rather than upon that of the department. Again, contrary to the views of the Hawthorne researchers and of most subsequent interpreters, some empirical evidence is provided here supporting the positive influence of small group incentives upon weekly output rates in the first relay experiment and upon hourly output rates in the second relay experiment (fn. 9). However, the effects of incentives in the first relay experiment are minor relative to the three factors discussed above, and thus will not be considered further.

In this statistical analysis of the first relay experiment, only the relatively small but consistent effect of rest pauses upon production quantity provides support for the contention that economic benefits result from humanitarian activity. The lack of substantial unexplained variance in any of the final models for output quantity indicates that the unmeasured supervisory

and social interaction variables were not very important economically. As Carey (1967) suggested in his incisive but non-quantitative critique, reevaluation of the experiments does not support the conclusions of the Hawthorne investigators.¹² Still, there remains Carey's (1967:403) question of how it was possible for "conclusions so little supported by evidence to gain so influential and respected a place within scientific disciplines and to hold this place for so long." One explanation for this enthusiastic embrace of something scientifically unproved may lie in the particular emphasis of the Hawthorne conclusions, another in the nonsubstantive nature of most criticism of them. Conclusions of the Hawthorne studies seem to have been congenial to persons who were in agreement with the prevailing economic system, but were prepared to proceed from simple materialistic notions about work motivation on to more complex social theories, which could be seen as more useful, humane, and democratic. Authors who appear to have interpreted the Hawthorne studies in this way include those in the volume edited by Cass and Zimmer (1975), and DeNood (1941), Friedman (1946), Homans (1941; 1949; 1950), Landsberger (1958), Miller and Form (1951), Nieder (1975), Sanford (1973), Shepard (1971), and Vroom (1964).

Most criticism in early years was ideological rather than substantive, in part directed, as noted by Landsberger (1958), against the ideology of Mayo (1919; 1933; 1945; 1947) and Whitehead (1936), and not particularly concerned with what the Hawthorne studies themselves had to say. This criticism did not treat seriously the main body of work by Roethlisberger and Dickson (1939), supplemented by Whitehead (1938). Examples of such

rather misdirected interpretations include the writings of Bell (1947), Gilson (1940), Lynd (1937), Bendix and Fisher (1949), and Schneider (1950). But complaints by Sheppard (1949; 1950) and Hampden-Turner (1970) of reactionary tendencies at Hawthorne have been given some plausibility by Homans's (1941) and Wilensky and Wilensky's (1951) observations that union activities failed at Western Electric. Other social scientists have been diverted by the Hawthorne effect, described by Roethlisberger (1941:14): "... If a human being is being experimented upon, he is likely to know it. Therefore, his attitudes toward the experiment and toward the experimenters become very important factors in determining his responses to the situation" (cf. also Dickson and Roethlisberger, 1966, and Bishop and Hall, 1971). This concept of influence upon an experiment through the experiment itself was found either erroneous or misleading by Cook and Campbell (1976), Katz and Kahn (1966), Parsons (1974), and Rubeck (1975). Sommer's (1968) conclusion, that the "errors" called placebo or Hawthorne effect need themselves to be evaluated and understood, is most pertinent.

Perhaps discouraged by the inaccessibility of numerical data from the experiments (although outputs were graphed and most independent variables described by Whitehead, 1938), not one of the numerous commentators has attempted a quantitative interpretation of changes in output rates for nearly 50 years.¹³ Indeed, except for Hare (1967) and Parsons (1974), most of the interpreters of the first relay experiment appear not to have recognized that there were more than 13 experimental periods, even though 15 periods had been described by Pennock (1930) and a total of 24 by Whitehead (1938; cf. fn. 6).

In the social sciences—particularly where complex beliefs and processes are

¹² Acker and Van Houten (1974:156) similarly concluded that "the cumulative effect of coercion, paternalistic treatment, and special rewards resulted in a rise in productivity." Others that criticized the early evaluations of the first relay experiment are Argyle (1953), Blum and Naylor (1968), Farris (1969), Locke (1976), Moore (1947), Sykes (1965), and Viteles (1953). Acker and Van Houten further suggested that results for the female workers in the first relay experiment might not be applicable to male workers.

¹³ Parsons (1974) offered a behavioral theory to replace the social interaction theory of the Hawthorne investigators, but this theory also is rendered implausible by the present analyses. Still, quantitative testing of the feedback mechanism suggested by Parsons should be possible using the continuous production record available from the first relay experiment.

involved—there seems to be no substitute for quantitative analysis. Whether there are only few data available or, as in the present case, where there exists a massive body of data and description, quantitative analysis enables the scientist to separate fact from fiction. Much of the information from the Hawthorne experiments remains to be tapped and interpreted with this aim.¹⁴

CONCLUSION

This first statistical interpretation of the major Hawthorne experiment leads to conclusions different from those heretofore drawn. Most of the variance in production rates during the first relay experiment could be explained by measured variables. To assume that output changes resulted from unmeasured changes in the human relations of workers therefore seems injudicious, even though it was the assumption of the Hawthorne researchers and has been accepted and built upon by many social scientists over the past several decades.

The Hawthorne experiments, most of which involved small groups of workers, are exceptional in the accumulation of information over extensive periods of time under actual working conditions. The experiments drew attention to small group processes, and the studies' conclusions led to widespread acceptance of human relations as a primary factor in worker performance. Following dissemination of the findings, previously accepted and conceptually simpler mechanisms such as those of scientific management (Taylor, 1911) tended to be given less emphasis as determinants of work performance. These

include the possible benefits of fatigue reduction, use of economic incentives, the exercise of discipline, and other aspects of managerial control. But it is precisely such factors to which we are directed by empirical analyses of the Hawthorne data. In particular, the discharge and replacement of two somewhat insubordinate workers were followed by higher group and individual production rates in the first relay experiment. Fairly strong evidence has been provided in recent years showing that proclivity to exert close managerial control can benefit the economic performance of individual managers (Miner, 1965), of organizations (Kock, 1965), and of whole societies (Franke, 1973; 1974; 1977). If the empirical results from the Hawthorne experiments and from these more recent studies contain some general applicability to economic organizations, then more of our attention as social scientists might well be directed to managerial characteristics and processes and somewhat less to the human relations of workers. Quantitative analyses of the data from Hawthorne, as well as empirical studies of work groups in the decades subsequent (cf. Stogdill, 1974), unfortunately do not support a contention that improvements in human relations lead to improved economic performance. On the other hand, such activities as participative management, industrial democracy, and sensitivity or consideration training may have benefits transcending the criteria considered here.

The analytical procedures employed in the present study suggest feasibility of examining closely the building blocks of our disciplines, especially when quantitative information is available. This has long been done in the physical sciences, where development routinely includes the process of critical scientific review, secondary analysis, and replication of important studies. There appears great need as well as opportunity for such activities in the social sciences.

¹⁴ A "Guide to Hawthorne Records," which provides entry to the UWM and WPI microfilm files and to the comprehensive index of Mallach and Smith (1977), may be obtained from Franke. All data used in the present analysis are provided in Appendices 1 and 2. We wish to encourage reappraisal of our calculations, as well as further investigation of these historically important experiments.

APPENDIX 1

Group Data from First Relay Experiment *

Period Number (Dates Included)	(1) Hourly Output	(2) Weekly Output	(3) Repair Time	(4) Hours Per Day	(5) Days Per Week	(6) Net Hours Per Week	(7) Weeks Per Period	(8) Sched- uled Rest Stops	(9) Sched- uled Rest Time	(10) Volun- tary Rest Time
1 (4/25-5/10/27)	49.7	2385.60	—	8.75	5.5	48.00	2	0	0	—
2 (5/10-6/11/27)	49.1	2356.80	—	8.75	5.5	48.00	5	0	0	10.5
3 (6/13-8/6/27)	51.0	2448.00	14.9	8.75	5.5	48.00	8	0	0	13.7
4 (8/8-9/10/27)	52.1	2452.87	18.5	8.75	5.5	47.08	5	2	10	9.0
5 (9/12-10/8/27)	55.1	2543.97	26.4	8.75	5.5	46.17	4	2	20	9.5
6 (10/10-11/5/27)	55.6	2515.90	31.7	8.75	5.5	45.25	4	6	30	0.5
7 (11/7/27-1/21/28)	55.9	2552.95	18.8	8.75	5.5	45.67	11	2	25	8.4
8 (1/23-3/10/28)	61.9	2692.22	23.2	8.25	5.5	43.17	7	2	25	2.8
9 (3/12-4/7/28)	63.9	2598.81	17.2	7.75	5.5	40.67	4	2	25	2.3
10 (4/9-6/30/28)	61.8	2822.41	15.8	8.75	5.5	45.67	12	2	25	5.5
11 (7/2-9/1/28)	62.8	2616.88	19.4	8.75	5.0	41.67	9	2	25	6.4
12 (9/3-11/24/28)	60.7	2913.60	13.4	8.75	5.5	48.00	12	0	0	14.3
13 (11/26/28-6/29/29)	66.5	3039.06	14.4	8.75	5.5	45.67	31	2	25	7.0
14 (7/1-8/31/29)	63.3	2637.71	48.5	8.75	5.0	41.67	9	2	25	6.9
15 (9/2/29-4/5/30)	66.2	3023.35	40.2	8.75	5.5	45.67	31	2	25	5.2
16 (4/7-5/3/30)	69.7	3183.20	30.5	8.75	5.5	45.67	4	2	25	4.6
17 (5/5-10/25/30)	69.2	2624.06	22.0	8.00	5.0	37.92	25	2	25	5.4
18 (10/29/30-2/7/31)	69.6	2406.79	22.0	8.00	4.5	34.58	15	2	25	7.1
19 (2/9-5/23/31)	69.3	2396.39	26.9	8.00	4.5	34.58	15	2	25	7.0
20 (5/25-11/14/31)	68.6	2601.31	24.2	8.00	5.0	37.92	25	2	25	6.0
21 (11/16-12/5/31)	69.6	2110.97	—	8.00	4.0	30.33	3	2	25	5.2
22 (12/7/31-2/6/32)	71.7	2718.86	—	8.00	5.0	37.92	9	2	25	4.6
23 (2/8-2/27/32)	71.5	2168.60	—	8.00	4.0	30.33	3	2	25	—

* For specific microfilm sources, contact first author (cf. fn. 14). Values of categoric variables: X_{11} over periods 1-7: 0, periods 8-23: 1; X_{12} over periods 1-14: 0, period 15: 0.74, periods 16-23: 1; X_{13} over periods 1-13 and 16-23: 0, periods 14-15: 1; X_{14} over periods 1-13 and 18-23: 0, period 14: 0.44, periods 15-16: 1, period 17: 0.04; X_{15} over periods 1-2: 0, periods 3-23: 1.

APPENDIX 2
Individual Data from First Relay Experiment*

Operator:	(1) Hourly Output, units/hour					(3) Repair Time, min./day					(10) Voluntary Rest Time, min./day				
	1A+1	2A+2	3	4	5	1A+1	2A+2	3	4	5	1A+1	2A+2	3	4	5
Period															
1	50.5	49.7	49.7	49.7	48.3	—	—	—	—	—	—	—	—	—	—
2	47.8	48.0	49.5	51.1	48.9	—	—	—	—	—	10.0	12.6	12.0	12.7	5.3
3	48.4	50.4	53.6	52.2	50.5	19.0	13.5	12.3	10.6	19.0	13.7	17.0	13.0	16.5	8.3
4	51.5	50.7	53.6	53.6	50.8	19.7	16.3	13.4	9.4	32.8	8.8	4.8	11.0	13.9	6.5
5	54.1	55.4	56.9	56.1	52.9	34.6	37.4	17.4	11.4	31.5	9.8	9.3	9.3	12.8	6.5
6	55.2	54.7	56.8	56.8	54.6	38.0	36.3	21.2	12.5	50.3	0.3	1.1	0.2	0.5	0.6
7	54.0	53.9	58.9	58.2	54.2	23.5	23.0	14.2	6.2	27.0	11.4	6.1	10.3	9.0	5.1
8	62.8	64.5	62.2	63.1	56.8	39.0	25.2	19.0	6.3	26.4	0.5	5.3	1.3	4.5	2.2
9	65.5	68.0	63.0	63.5	59.5	30.5	16.8	16.3	6.0	16.3	—	4.6	0.5	5.0	1.6
10	63.9	64.9	62.1	62.8	55.2	24.1	20.0	12.5	3.5	18.7	0.2	7.6	3.6	11.2	4.9
11	65.6	66.4	63.9	62.9	55.0	33.6	27.6	13.1	9.6	13.2	—	8.9	4.3	13.8	4.8
12	62.5	63.9	59.7	61.3	56.1	23.7	22.9	11.0	3.8	5.6	15.0	20.6	11.2	15.8	8.9
13	67.4	71.9	64.3	69.1	59.7	22.3	19.0	13.8	7.8	8.9	5.0	10.0	5.4	9.2	5.6
14	64.5	70.6	62.0	69.4	50.0	59.9	42.7	40.7	30.0	69.0	6.1	8.4	7.1	8.0	5.1
15	68.8	73.9	64.0	71.3	52.7	48.8	36.7	36.8	18.4	60.2	5.1	6.0	5.1	5.2	4.5
16	72.7	76.8	67.5	73.0	59.6	39.5	31.7	31.7	19.2	30.5	3.9	4.4	6.1	4.8	4.2
17	69.6	75.9	65.6	73.7	61.5	28.0	20.3	26.1	14.6	21.1	4.6	5.0	6.3	4.9	6.4
18	69.1	76.4	64.9	72.8	64.9	21.0	25.0	26.0	13.7	24.5	6.3	5.7	9.2	6.2	8.0
19	72.9	74.8	64.1	69.6	65.3	32.8	30.0	30.6	19.6	21.3	6.1	5.9	8.3	6.9	7.6
20	72.5	76.3	63.4	66.5	64.0	28.7	25.0	26.0	22.7	18.6	5.0	5.4	6.8	6.6	5.8
21	70.2	72.5	67.0	68.9	67.7	—	—	—	—	—	3.7	4.6	6.2	5.5	5.8
22	75.0	76.4	68.6	73.1	65.8	—	—	—	—	—	3.6	3.6	5.0	5.1	5.6
23	77.1	78.0	66.9	73.1	67.1	—	—	—	—	—	—	—	—	—	—

* For specific microfilm sources, contact first author (cf. fn. 14). Individual weekly output rates can be calculated by multiplying individual values of variable (1), above, by variable (6) from Appendix 1.

REFERENCES

- Acker, Joan and Donald R. Van Houten
1974 "Differential recruitment and control: the sex structuring of organizations." *Administrative Science Quarterly* 19:152-63.
- Argyle, Michael
1953 "The relay assembly test room in retrospect." *Occupational Psychology* 27:98-103.
- Bass, Bernard M.
1960 *Leadership, Psychology, and Organizational Behavior*. New York: Harper.
- Bass, Bernard M. and Gerald V. Barrett
1972 *Man, Work, and Organizations*. Boston: Allyn and Bacon.
- Beer, Michael
1976 "The technology of organizational development." Pp. 937-93 in M. D. Dunnette (ed.), *Handbook of Industrial and Organizational Psychology*. Chicago: Rand McNally.
- Bell, Daniel
1947 "Adjusting men to machines." *Commentary* 3:79-88.
- Bendix, Reinhard and Lloyd H. Fisher
1949 "The perspectives of Elton Mayo." *Review of Economics and Statistics* 31:312-9.
- Bishop, Ronald C. and James W. Hill
1971 "Effects of job enlargement and job change on contiguous but nonmanipulated jobs as a function of workers' status." *Journal of Applied Psychology* 55:175-81.
- Blalock, Hubert M., Jr.
1961 *Causal Inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.
- Blum, Milton L. and James C. Naylor
1968 *Industrial Psychology: Its Theoretical and Social Foundations*. New York: Harper and Row.
- Bradford, L. P., J. Gibb and K. Benne (eds.)
1964 *T-group Theory and Laboratory Method*. New York: Wiley.
- Campbell, Donald T. and Julian C. Stanley
1963 "Experimental and quasi-experimental designs for research on teaching." Pp. 171-246 in N. L. Gage (ed.), *Handbook of Research on Teaching*. Chicago: Rand McNally.
- Carey, Alex
1967 "The Hawthorne studies: a radical criticism." *American Sociological Review* 32:403-16.
- Cass, Eugene Louis and Frederick G. Zimmer (eds.)
1975 *Man and Work in Society*. New York: Van Nostrand Reinhold.
- Cherrington, David J., H. Joseph Reitz, and William E. Scott, Jr.
1971 "Effects of contingent and non-contingent reward on the relationship between satisfaction and task performance." *Journal of Applied Psychology* 55:531-6.
- Cook, Thomas D. and Donald T. Campbell
1976 "The design and conduct of quasi-experiments and true experiments in field settings." Pp. 223-326 in M. D. Dunnette (ed.), *Handbook of Industrial and Organizational Psychology*. Chicago: Rand McNally.
- DeNood, Neal B.
1941 "Book review: management and the worker." *American Sociological Review* 6:304-5.
- Dickson, William J. and F. J. Roethlisberger
1966 *Counseling in an Organization*. Boston: Harvard University.
- Dogan, M. and S. Rokkan (eds.)
1969 *Quantitative Ecological Analysis in the Social Sciences*. Cambridge, Ma.: MIT Press.
- Durbin, J. and G.S. Watson
1950 "Testing for serial correlation in least squares regression, I." *Biometrika* 37:409-28.
- 1951 "Testing for serial correlation in least squares regression, II." *Biometrika* 38:159-78.
- Elliott, J. Walter
1973 *Economic Analysis for Management Decisions*. Homewood: Irwin.
- Farris, George F.
1969 "The drunkard's search in behavioral science." *Personnel Administration* 32:10-8.
- Fleishman, E. A., E. F. Harris, and H. E. Burt
1955 *Leadership and Supervision in Industry*. Columbus: Bureau of Educational Research, Ohio State University.
- Franke, Richard Herbert
1973 "Critical factors in the post-war economic growth of nations: review of empirical studies and implications for participative organization." Pp. 107-19 in E. Pusić (ed.), *Participation and Self Management*, Vol. 5. Zagreb: University of Zagreb.
- 1974 *An Empirical Appraisal of the Achievement Motivation Model Applied to Nations*. Ph.D. dissertation, Graduate School of Management, University of Rochester.
- 1977 "Culture and industrial development." Paper presented at the International Conference on Social Change and Organizational Development, Inter-University Centre of Post-Graduate Studies, Dubrovnik.
- Friedman, Georges
[1946] *Problèmes Humains du Machinisme Industriel*. Paris: Librairie Gallimard. Trans. and ed. by Harold L. Sheppard as *Industrial Society: the Emergence of the Human Problems of Automation*. Glencoe: Free Press.
- Galtung, Johan
1967 *Theory and Methods of Social Research*. New York: Columbia University Press.
- Gilson, Mary B.
1940 "Book review: management and the worker." *American Journal of Sociology* 46:98-101.
- Golembiewski, Robert T. and Arthur Blumberg
1970 *Sensitivity Training and the Laboratory Approach*. Itasca: Peacock.
- Hampden-Turner, Charles
1970 *Radical Man*. Cambridge, Ma.: Schenkman.

- Hare, A. Paul
1967 "Small group development in the relay assembly testroom." *Sociological Inquiry* 37:169-82.
- Homans, George Caspar
1941 "The Western Electric researches." Pp. 56-99 in National Research Council Committee on Work in Industry, *Fatigue of Workers: Its Relation to Industrial Production*. New York: Reinhold.
1949 "Some corrections." *Review of Economics and Statistics* 31:319-21.
1959 *The Human Group*. New York: Harcourt Brace.
- Johnston, J.
1963 *Econometric Methods*. New York: McGraw-Hill.
- Kahn, Robert L.
1975 "In search of the Hawthorne effect." Pp. 49-63 in Eugene Louis Cass and Frederick G. Zimmer (eds.), *Man and Work in Society*. New York: Van Nostrand Reinhold.
- Katz, Daniel and Robert L. Kahn
1966 *The Social Psychology of Organizations*. New York: Wiley.
- Kock, Sven E.
1965 *Företagsledning och motivation*. (Management and Motivation, with English summary.) Helsinki: Affärssekonomisk Förlagsförening.
- Landsberger, Henry A.
1958 *Hawthorne Revisited*. Ithaca: Cornell University.
- Likert, Renis
1967 *The Human Organization*. New York: McGraw-Hill.
- Locke, Edwin A.
1976 "The nature and causes of job satisfaction." Pp. 1297-349 in Marvin D. Dunnette (ed.), *Handbook of Industrial and Organizational Psychology*. Chicago: Rand McNally.
- Lynd, Robert S.
1937 "Book review: leadership in a free society." *Political Science Quarterly* 52:590-2.
- Mallach, Stanley and Steven Smith
1977 *Records of the Industrial Relations Experiment Carried Out by the Western Electric Company at the Hawthorne Works, Hawthorne, Illinois*. Index, University of Wisconsin-Milwaukee Library, Milwaukee.
- Marrow, Alfred J.
1975 "Management by participation." Pp. 33-48 in Eugene Louis Cass and Frederick G. Zimmer (eds.), *Man and Work in Society*. New York: Van Nostrand Reinhold.
- Mayo, Elton
1919 *Democracy and Freedom*. Melbourne: Macmillan.
1933 *The Human Problems of an Industrial Civilization*. New York: Viking.
1945 *The Social Problems of an Industrial Civilization*. Boston: Harvard University.
1947 *The Political Problems of an Industrial Civilization*. Boston: Harvard University.
- Miller, Delbert C. and William H. Form
1951 *Industrial Society*. New York: Harper.
- Miner, John B.
1965 *Studies in Management Education*. New York: Springer.
- Moore, Wilbert E.
1947 "Current issues in industrial sociology." *American Sociological Review* 12:651-7.
- Nieder, Peter
1975 "Zum zusammenhang zwischen führungsverhalten, produktionsniveau und zufriedenheit." *Gruppendynamik* 2:127-39.
- Parsons, H. M.
1974 "What happened at Hawthorne?" *Science* 183:922-32.
- Pennock, G. A.
1930 "Industrial research at Hawthorne: an experimental investigation of rest periods, working conditions and other influences." *Personnel Journal* 8:296-313.
- Price, James L.
1968 *Organizational Effectiveness: An Inventory of Propositions*. Homewood: Irwin.
- Pusić, E. (ed.)
1973 *Participation and Self Management*. 6 Vols. Zagreb: University of Zagreb.
- Robinson, W. S.
1959 "Ecological correlations and the behavior of individuals." *American Sociological Review* 15:352-7.
- Roethlisberger, F. J.
1941 *Management and Morale*. Cambridge, Ma.: Harvard University Press.
- Roethlisberger, F. J. and William J. Dickson
1939 *Management and the Worker*. Cambridge, Ma.: Harvard University Press.
- Rubeck, Patricia A.
1975 "Hawthorne concept—does it affect reading progress?" *The Reading Teacher* 28:375-9.
- Sales, S. M.
1966 "Supervisory style and productivity: review and theory." *Personnel Psychology* 19:275-86.
- Sanford, Aubrey C.
1973 *Human Relations: Theory and Practice*. Columbus: Merrill.
- Schneider, Louis
1950 "An industrial sociology—for what ends?" *Antioch Review* 10:407-17.
- Shepard, Jon M.
1971 "On Alex Carey's radical criticism of the Hawthorne studies." *Academy of Management Journal* 14:23-32.
- Sheppard, Harold L.
1949 "The treatment of unionism in 'managerial sociology'." *American Sociological Review* 14:310-3.
1950 "The social and historical philosophy of Elton Mayo." *Antioch Review* 10:396-405.
- Sommer, Robert
1968 "Hawthorne dogma." *Psychological Bulletin* 70:592-5.
- Stogdill, Ralph M.
1974 *Handbook of Leadership*. Glencoe: Free Press.
- Sykes, A. J. M.
1965 "Economic interest and the Hawthorne researches." *Human Relations* 18:253-63.

- Taylor, Frederick W.
1911 *The Principles of Scientific Management*. New York: Harper.
- Theil, H. and A. L. Nager
1961 "Testing the independence of regression disturbances." *American Statistical Association Journal* 56:793-806.
- Thorndike, E. L.
1939 "On the fallacy of imputing the correlations found for groups to the individuals or smaller groups composing them." *American Journal of Psychology* 52:122-4.
- Viteles, Morris S.
1953 *Motivation and Morale in Industry*. New York: Norton.
- Vroom, Victor H.
1960 *Some Personality Determinants of the Effects of Participation*. Englewood Cliffs: Prentice-Hall.
- 1964 *Work and Motivation*. New York: Wiley.
- Whitehead, T. N.
1936 *Leadership in a Free Society*. Cambridge, Ma.: Harvard University Press.
- 1938 *The Industrial Worker*. 2 Vols. Cambridge, Ma.: Harvard University Press.
- Whyte, William F.
1955 *Money and Motivation: An Analysis of Incentives in Industry*. New York: Harper and Row.
- Wilensky, Jeanne L. and Harold L. Wilensky
1951 "Personnel counseling: the Hawthorne case." *American Journal of Sociology* 57:265-80.

THE MYTH OF SOCIAL CLASS AND CRIMINALITY: AN EMPIRICAL ASSESSMENT OF THE EMPIRICAL EVIDENCE*

CHARLES R. TITTLE
Florida Atlantic University

WAYNE J. VILLEMEZ
University of Illinois, Chicago Circle

DOUGLAS A. SMITH
Indiana University

American Sociological Review 1978, Vol. 43 (October):643-656

Thirty-five studies examining the relationship between social class and crime/delinquency are reduced to comparable statistics using instances where the relationship was studied for specific categories of age, sex, race, place of residence, data type, or offense as units of analysis. The findings from 363 instances are summarized and patterns are identified. The overall results show only a slight negative relationship between class and criminality, with self-report studies reflecting lower associations than official statistics studies. Moreover, analysis demonstrates a clear historical decline in magnitude of association to the point where both self-report and official statistics studies done in the current decade find no class variation. This historical trend is shown to be due to changes in the findings from studies using official statistics as indicators of criminality. Alternative interpretations are discussed, but all lead to serious doubts about the adequacy of theories of deviance that contain assumptions of class differences.

Social scientists long have assumed an intimate linkage between a variety of social pathologies and injustice or inequity in the distribution of societal resources. This is a reasonable assumption because differences in social power and advantage imply differences across the whole range of life chances. But a relationship between the distribution of social resources and behavioral manifestations is more easily

justified on theoretical than empirical grounds. For one thing, concentration of resources into distinguishable categories never has been measured clearly enough to permit firm conclusions about relationships. Indeed, controversy about the extent of resource concentration has pervaded the stratification literature. At one point social class was a widely accepted concept for describing such concentrations, but following a concerted attack in the late fifties and early sixties (e.g., Cutright, 1968; Glenn, 1967; Laumann, 1966; Ossowski, 1963; Nisbet, 1959; Rose,

* Address all communications to: Charles R. Tittle; Department of Sociology/Social Psychology; Florida Atlantic University; Boca Raton, FL 33431.

1958), the focus of research changed from classes to strata. And since stratum implies more ambiguous boundaries than does class, the hypothesis of behavioral differences among those in various categories of resource concentration became even more difficult to test. As a result, interest in the question diminished among students of stratification (cf. Glenn et al., 1970).

But in some subdisciplines interest in the purported relationship has continued to flourish. This is especially true for the field of deviance/criminology. Social class (however defined) has always been a central variable in studies of crime and delinquency, and has been important in almost every theory. Although several patterns of relationship between class and crime have been hypothesized (Hirschi, 1972) the most popular theories predict an inverse relationship between social position and criminality (Cohen, 1955; Gove, 1975; Merton, 1968; Miller, 1958). And despite some critical assessments (Box and Ford, 1971; Doleschal, 1970; Empey, 1967), a belief persists that the bulk of evidence actually demonstrates such a relationship (Bytheway and May, 1971; Cressey, 1966; Reiss, 1976; Rossides, 1976; Wheeler, 1966). So firm is this belief that at least one recent book has been written to account for the "diverse empirical findings that link social inequality and deviant behavior, particularly in American society" (Hewitt, 1970:3). Further, even though many acknowledge that self-report studies contradict the general hypothesis and are now careful to limit their generalizations to statements about social class and officially recorded crime or delinquency (Cohen and Short, 1971; Reid, 1976), confidence that at least that relationship has been unquestionably established is almost universal (Hood and Sparks, 1970; Reiss, 1976).

But despite general acceptance, there is good reason to question whether the evidence does in fact demonstrate that the social status of individuals is related inversely to criminal or delinquent behavior. For one thing, methodological limitations undermine the generalizability, applicability, or validity of much of the data. For example, many of the frequently

cited studies report relationships between class measures and crime for ecological areas rather than for individuals (Chilton, 1964; Lander, 1954; Shaw and McKay, 1969; Slatin, 1969; see also Gordon, 1967). Although there is legitimate sociological value in ecological analysis, it does not necessarily permit inferences on the individual level (Robinson, 1950). Lower status areas may have higher crime rates because a small proportion of people within those areas commit a lot of crimes or because outsiders come into these areas to do mischief. And, it is possible that people who live in low status areas and commit crimes are not necessarily themselves of low status, since the composition of urban units such as census tracts are often diverse. True enough, some observational data, victimization surveys, and official statistics concerning the characteristics of arrestees are consistent with an inference that ecological correlations reflect an underlying negative association between the social status of individuals and criminal behavior (Reiss, 1976), but many of these data are themselves questionable, and there is no direct evidence which demonstrates the individual-area connection. In fact, Johnstone (1977) reports data from a sample of youth in the Chicago SMSA that show a great deal of inconsistency between contextual status variation in self-report delinquency and variations by family status. Moreover, ecological correlations between crime and social status have been interpreted as self-fulfilling prophecies related to police deployment patterns and as functions of the greater visibility of criminal behavior by those who live in lower status areas (Chambliss, 1975:135). Hence, in the absence of direct evidence linking the ecological findings to individual characteristics and in view of contrary interpretations of the area correlations, the ecological evidence has to be regarded as problematic.

In addition, several ostensibly relevant investigations have not in fact employed indicators of individual status (Cohen, 1969; Conger and Miller, 1966; Garrett and Short, 1975; Meade, 1973; Wolfgang et al., 1972; Won and Yamamoto, 1968). For example, the massive work by

Wolfgang and his associates used census tract median income as a measure of family status rather than actual family income. Although tract data do allow investigation of contextual effects, there is good reason to question whether they are valid proxies for individual status characteristics. After all, census tracts are usually quite heterogeneous internally. The 1970 census reveals that 75% of all the census tracts in Philadelphia include families with incomes ranging from less than \$1,000 to over \$25,000. Thus there easily could be dramatic differences between the actual family income of subjects and the median income of the tract in which that family resides. As a case in point, only 52% of a sample of Chicago youth surveyed in 1972 had family status characteristics that were even in the same third of a distribution as the status characteristics of the census tracts in which they resided (Johnstone, 1977).

Further, some of the research is weakened by crude measurement of either social class (Hardt and Peterson, 1968; Polk et al., 1974) or crime/delinquency (Stinchcombe, 1964), and in several instances analysis has been primitive or attenuated (Conger and Miller, 1966; Winslow, 1967; Won and Yamamoto, 1968).

But apart from methodological questions, there is a second reason the evidence is unconvincing. That is simply that it is less voluminous and comprehensive than usually is thought. Although at least 40 studies of the class/crime relationship exist in the literature, there are glaring deficiencies in their representativeness. Many studies have been limited to only one racial/sexual category (Empey and Erickson, 1966; Reiss and Rhodes, 1961), and samples often have been less than comprehensive, even for the specific target population. About 85% of all the studies concern juvenile misconduct, and most focus only on white males. And despite frequent references (without citation) in the literature to "many studies," we are able to find only 16 investigations that used official police contact or court delinquency figures and only seven studies examining official arrest or conviction data for adults.

Finally, since the various studies use different methodologies, types of data, forms of reporting, samples, and statistics, it is almost impossible to draw meaningful general conclusions about the nature of the class/crime relationship that might be reflected in these 40 studies. For example, Tittle and Villemez (1977) prefaced their recent study with an attempt to summarize the state of knowledge by counting the proportion of studies that reported a general negative relationship between SES and crime/delinquency. But the weakness of this approach is obvious. It is extremely inefficient and likely to product erroneous conclusions because it counts each study equally regardless of the nature and size of the sample, or of the magnitude of association revealed in each instance. But almost any other method would be equally inefficient since the studies are simply not comparable in their original form. The fact is, we do not know what accumulated research tells us about the relationship between social class and crime/delinquency because the available evidence has never been sorted out adequately.

The purpose of this paper is to bring order out of the existing chaos in two ways. First, we attempt to reduce the extant data to a single comparable base. And second, we treat the measures of association derived from the existing literature as dependent variables, and we attempt to account for variations in them by using characteristics of the various studies as independent variables. This approach allows the plethora of investigations to be weighed against one another in a meaningful way and permits nonpolemical assessment of the general findings as well as an analysis of trends in results.

METHODS

We attempted to identify every instance in the literature where the relationship between a measure of individual class position and crime/delinquency indicators has been reported (although we do not guarantee that we might not have missed something). Each report of that relationship for a specific category of age, sex, race, place of residence, data type, or offense was

treated as a separate "instance" and where no differentiation by these variables was made, the overall relationship reported in the study was considered one "instance." For each instance, data reported in the original study were rearranged into ordinal contingency tables and gamma measures of association were calculated. By this method we were able to reduce the evidence to 363 separate gammas that serve as independent indicators of the nature of the relationship between social class and deviance. Of course, some of the studies that exist in the literature do not report complete enough data to permit derivation of contingency tables, and some of the reported data are inappropriate for contingency table analysis. For example, some studies report only the measure of association or a statistical test without raw data (Cohen, 1969; Kvaraceus, 1944; Meade, 1973). Where the reported measure is something besides a gamma and the data are not reconstructable, we were unable to use the material in our scheme. Moreover, some studies analyse such things as variations in official rates of one kind or another, but in the absence of knowledge of the actual number of people in each SES category in the particular population, it is impossible to fill in the cells of a contingency table (Bonger, 1916; Clark and Wenninger, 1962; Conger and Miller, 1966; Garrett and Short, 1975). One study reports only the percent of offenses committed by various status categories without raw data for individuals (Empey and Erickson, 1966). And one contemporary self-report study, based on a comprehensive sample of youth in the Chicago area, reports only mean delinquency scores for categories of social status (Johnstone, 1977). Nevertheless, the 363 gammas we were able to derive represent accumulated evidence from 35 separate studies.¹

¹ There were a total of nine studies (cited above) unusable for technical reasons. The 35 used include: Akers, 1964; Arnold, 1965; Belson, 1975; Berger and Simon, 1974; Cameron, 1964; Casparis and Vaz, 1974; Christie et al., 1965; Dentler and Monroe, 1961; Douglas et al., 1966; Elliott, 1962; Erickson, 1973; Gould, 1969; Gold, 1966; Green, 1970; Hardt and Peterson, 1968; Havighurst et al., 1962; Hirschi, 1969; Hollingshead, 1947; Kelly, 1975; Kelly and

This compilation of studies excludes those which report data representing ecological units only, because we believe that the connection between ecological level data and individual characteristics is too problematic to justify their inclusion. As noted in the introduction, it is easy to make an incorrect inference; there is no supplementary evidence which directly shows a connection between the two levels for the crime-delinquency/class question, and there are contradictory explanations of the meaning of area correlations. Moreover, heterogeneity of individual and status characteristics within most census tracts causes us to be even more skeptical of the ability to draw conclusions about individuals from that type of study. But even if it had been reasonable to consider the ecological studies, we would have found it impossible to reduce the data to a form that would permit gamma statistics comparable with the other studies included.

The set of usable studies does, however, include the Wolfgang et al. (1972) and Won and Yamamoto (1968) studies in which individual measures of social class were derived from census tract characteristics. In the beginning we intended to include all studies where there was an individual indicator of class, even if that indicator was estimated from aggregated data. But four of six such studies turned out to be unusable for one or another of the technical reasons stated before. Although we believe the Wolfgang type evidence is likely to be misleading, we considered those studies because they are nevertheless part of the collective pool of information from which inferences about the class/crime-delinquency relationship on an individual level can be drawn, in the same way that studies which suffer from poor sampling or crude analysis are.

Pink 1975; McDonald, 1969; Nye et al., 1958; Polk et al., 1974; Reiss and Rhodes, 1961; Stinchcombe, 1964; Tittle and Villemez, 1977; Voss, 1966; Walberg et al., 1974; Warner and Lunt, 1941; Williams and Gold, 1972; Winslow, 1967; Wolf, 1962; Wolfgang et al., 1972; Won and Yamamoto, 1968; and part of the data from Empey and Lubeck, 1971. The Utah data reported by Empey and Lubeck appear to be the same as those analysed by Erickson (1973). Therefore we extract from Empey and Lubeck only those data gathered in Los Angeles.

Second, in an effort to explain variations in the magnitude and sign of the different gammas, we performed a multivariate analysis using the following as control or independent variables: (1) the type of sample; (2) the size of the sample; (3) the type of place where the study was conducted (rural, urban, or combination); (4) the size of the area from which the sample was drawn; (5) the area of the world or area of the country (for studies done in the U. S.) where the study was conducted; (6) the age, sex, and race of the persons in the given situation; (7) the type of data (self-reported, official, or combination); (8) the year the data were gathered; (9) the number of statuses differentiated in the measure of SES; (10) sophistication of the SES measure, as indicated by the number of dimensions considered and the method by which the data were combined; (11) the specific indicator or indicators of status used (occupation, education, income, other, or specific combinations); (12) the range of scores on the crime/delinquency indicator; (13) the kind of measure of crime/delinquency (incidence, frequency, or seriousness); (14) the sophistication of the measure of crime/delinquency; and (15) the type of deviance (major and minor youth offenses, personal offenses, property crimes, violent crimes, and undifferentiated official records).

FINDINGS

The basic findings are presented in Table 1, which shows average gammas for various categories of age, sex, and race. Contrary to general theoretical expectations and widespread popular opinion, the data as a whole show only a very slight negative relationship between social class and crime/delinquency. The overall gamma for the 363 instances (fourth panel) is only $-.09$, a figure which indicates almost no relationship. Indeed, a gamma of this magnitude could result from a mere two or three point difference between the percent of upper and lower class individuals displaying criminal tendencies. Examining column totals in panel four, one can see that the relationship is similarly weak in instances where only males are in-

Table 1. Mean Class/Criminality Gammas for Studies of Various Age, Race, and Sex Subgroups

	Youth			Adults			Youth and Adults			All Studies		
	Male		Total	Male		Total	Male		Total	Male		Total
	Female	Both		Female	Both		Female	Both		Female	Both	
White (N)	-.15 (38)	-.17 (3)	-.11 (63)	+.03 (19)	-.19 (12)	-.06 (31)	+.05 (18)	-.09 (18)	-.02 (36)	-.06 (75)	-.13 (33)	-.07 (130)
Nonwhite (N)	-.29 (3)	-.14 (2)	-.10 (24)	— (0)	— (0)	— (0)	+.37 (6)	-.03 (6)	+.17 (12)	+.15 (9)	-.06 (8)	-.01 (36)
Both (N)	-.11 (113)	-.09 (41)	-.09 (187)	— (0)	— (0)	-.46 (3)	-.63 (1)	-.54 (1)	-.67 (5)	-.11 (114)	-.10 (42)	-.12 (197)
Total (N)	-.12 (154)	-.10 (46)	-.10 (274)	+.03 (19)	-.19 (12)	-.09 (34)	+.10 (25)	-.10 (25)	-.06 (55)	-.08 (198)	-.11 (83)	-.09 (363)

cluded ($-.08$), where only females are included ($-.11$), and in mixed-sex instances ($-.10$). Row totals show the same for instances of whites ($-.07$), nonwhites ($-.01$), and of both combined ($-.12$). In short, the variance about our average gamma is small. With more thorough breakdowns by age, sex, and race of subjects (panels one through four), some larger average gammas result, but there is no consistent patterning. Although the signs of most gammas are negative (indicating some support for an inverse relationship between class and deviance), the measures themselves are usually quite small, and the signs are not all negative. The most stable results are in instances where young males are the object of study. Yet for this category, 154 individual instances yield a mean gamma of only $-.12$. There are categories with larger negative average gammas (all adult-only instances, $-.46$; all youth-adult mixed instances, $-.70$) but those averages are based on small N's (three and five, respectively) and are therefore not impressive given the variance found in subgroups with many instances. In fact, except for the two cases mentioned, if we set confidence limits about any of the mean gammas, those limits would include zero in every case. Thus, support for an overall negative class/criminality relationship is at best slight when the data are considered simply as a collectivity of evidence.

Although the average of the gammas is a

small negative figure, the individual instances show some marked differences in the magnitude of association. In an attempt to account for these differences, we examined the mean gammas for categories of the 15 independent variables specified previously in the methods section. With two exceptions, the results appear to be fairly stable under a wide range of conditions, with only minor and unreliable variations like those mentioned for age, sex, and race occurring among the categories of the 15 variables. For example, analysis of variations by type of offense shows little real difference (Table 2). The mean gamma for instances where youth offenses only are considered is $-.12$ ($N = 161$); where personal offenses are examined the mean gamma is $+.06$ ($N = 38$); for property crimes it is $-.06$ ($N = 105$); and it is $-.13$ ($N = 21$) for violent crimes. And where there are noticeable differences they disappear when the other variables are held constant. The two exceptions to this are the type of data and the decade in which the study was conducted.

Those differences clearly stand out. Table 3 reports the relevant data. Variations in the mean gamma from decade to decade show a steady decline in strength from the 1940s to the 1970s. The average gamma for the three instances where data were gathered prior to 1950 is $-.73$; as we move forward in time, it steadily diminishes. In the decade 1950-59 it is $-.31$; in the decade 1960-69 it is $-.13$; and for

Table 2. Mean Gammas for Categories of Type of Offense Studies by Decade in Which Study Was Done

Type of Offense	Before 1950	1950-59	1960-69	After 1970	N
Youth Offenses (\bar{X} gamma)	0 (—)	3 ($-.14$)	101 ($-.15$)	57 ($-.08$)	161 ($-.12$)
Personal Offenses (\bar{X} gamma)	0 (—)	0 (—)	0 (—)	38 ($+.06$)	38 ($+.06$)
Property Crime (\bar{X} gamma)	0 (—)	5 ($-.35$)	38 ($-.09$)	62 ($-.01$)	105 ($-.06$)
Violent Crime (\bar{X} gamma)	0 (—)	0 (—)	3 ($-.52$)	18 ($-.06$)	21 ($-.13$)
Undifferentiated Record (\bar{X} gamma)	3 ($-.73$)	2 ($-.47$)	12 ($-.17$)	0 (—)	17 ($-.30$)
Other (\bar{X} gamma)	0 (—)	0 (—)	16 ($-.11$)	5 ($-.04$)	21 ($-.09$)

Table 3. Mean Class/Criminality Gammas for Studies Using Self-Report and Official Data by Decade in Which the Study Was Done

Type of Data	Before 1950	1950-59	1960-69	After 1970	N
Self-Report	0	3	124	175	302
(\bar{X} gamma)	(—)	(— .04)	(— .11)	(— .03)	(— .06)
Official Statistics	3	7	46	5	61
(\bar{X} gamma)	(— .73)	(— .43)	(— .22)	(+ .04)	(— .25)
N	3	10	170	180	363
% Self-Report	0%	30%	73%	97%	85%
\bar{X} Gamma	— .73	— .31	— .13	— .03	— .09

those instances based on data gathered since 1970 the gamma is $-.03$. Moreover, as most scholars have suspected, instances based on officially recorded data tend to show much more marked negative associations than do those based on self-reports. The mean association for 302 instances based on self-report data is only $-.06$, but the mean for 61 instances using official data is $-.25$.

These findings also are derivable from another type of analysis. We regressed the gammas from all studies on a series of seven independent variables representing characteristics of the studies from which the gammas were drawn.² These seven were those which were found to be most salient in the earlier analysis. The resulting unstandardized equation is:

$$\begin{aligned}
 G = & .52 + .02ST - .00SS - .00AS \\
 & (.02) \quad (.00) \quad (.00) \\
 & - .10DT^* + .02YR^* \\
 & (.03) \quad (.00) \\
 & + .02NC + .16TT^* \\
 & (.01) \quad (.04)
 \end{aligned}$$

where * = double its standard error, and $R^2 = .26$, and ST = sample type, SS = sample size, AS = area size, DT = data type, YR = year of study, NC = number of social classes defined by the researcher, and TT = a time trend dummy differentiating those studies done prior to 1964 and

those done after 1963. The time trend variable was included because inspection of the data indicated that after 1963 the magnitude of the gammas began to decline markedly. The only significant coefficients found are those for type of data (official statistics vs. self-reports), year of study, and the trend dummy. Since data based on official statistics were assigned the higher number in the dichotomized code for type of data, the $-.10$ coefficient indicates that instances using official statistics tend to find negative class/crime relationships to a greater extent than do instances using self-report data. But because we are dealing with a dichotomous variable in which a unit change is the maximum possible, we can be even more specific. Given instances using two samples of exactly the same type and size, drawn from precisely the same size areas in a given year and employing the same number of defined social classes, the equation shows that we may expect the study employing official statistics to produce a gamma showing a .10 greater negative association for the same relationship than an instance employing self-report data.

The coefficient for the variable of the year of study indicates a trend toward a .02 per year decline in the size of negative gammas produced by all studies. But the time trend variable cannot be interpreted directly without further analysis. Following the technique employed by Masters (1975), we ran the regression: $G = a + b_1T_1 + b_2T_2 + b_3D$, where T_1 = the study years prior to 1964, numbered consecutively; T_2 = the study years including and subsequent to 1964, numbered consecutively; and D = a dummy variable coded

² The statistic gamma is not an inappropriate variable for regression analysis despite its clearly non-normal distribution. To have biased estimators it is necessary that $E(u) \neq 0$, and that condition has nothing to do with the natural distribution of gamma. Given our sample of studies and large N , we may safely assume the mean of the disturbance terms to be zero.

one if the data were gathered after 1963, and zero otherwise. The dummy variable allows independence of T_1 and T_2 , making significance-of-difference tests possible. Because the strongest interaction is between year of study and type of data (coefficients of a time/type multiplicative variable remained significantly greater than zero net of all mentioned variables), two separate time trend equations were run: one for official statistics instances ($N = 61$) and one for self-report instances ($N = 302$). Examining unstandardized coefficients for self-report studies, we find:

$$G = 1.01 - .01T_1 + .02T_2^* - .16D, R^2 = .04; \text{ and for official data: } G = .19 + .07T_1^* + .11T_2^* + .16D, R^2 = .56; \text{ where } * = \text{double its standard error.}$$

These equations reveal several interesting outcomes. First, the $T_2 - T_1$ difference for official data is over four times the magnitude of that difference for self-report data. Second, the coefficient of T_1 for self-report instances hardly differs from zero, indicating stability in the magnitude of the gammas from year to year up to 1964. And after that point there is but a slight decline in the size of the gammas. Thus, over the entire time period the gammas for self-report data have remained relatively constant. But gammas derived from official data show strong declines in both time periods with considerable escalation of the rate of decline after 1963 (an additional .04 per year). Although this $T_2 - T_1$ difference is not statistically significant ($t = 1.5, .10 > p > .05$), the time trend differences between self-report and official data gammas are significant (for T_1 , $t = 3.5, p < .001$; for T_2 , $t = 3.8, p < .001$). Finally, the time trend explains only 4% of the variance in gammas from self-report studies, while it alone explains 56% of the variance in gammas from official data.

It is interesting to note in passing that the only other coefficient of all those considered which approaches statistical significance is that for the number of classes that were differentiated by the researchers' measure. It produces a nonsignificant b , but the coefficient is of a size (relative to its standard error) to suggest

that, *ceteris paribus*, the use of a larger number of class categories in a study tends to lead to the discovery of a weaker class/criminality relationship. None of the other independent variables representing characteristics of the particular investigation shows a statistically significant net impact on the resultant gammas. Thus the only clearly important variables of those included are the type of data and the year of study. These two alone explain 19% of the variance in gammas, compared to 21% when all others (excluding the trend dummy) are added:

INTERPRETATION

Observing the large difference between the magnitude of the mean gamma revealed by self-report studies and that revealed by official statistics might tempt one to think that the observed historical decline in association between social class and crime/delinquency is an artifact of the greater use of self-report data in recent decades. But comparison of the results of the regression analyses for self-report and official statistics data as well as tabular analysis clearly disputes such an interpretation. There is a substantial year by year decline in the magnitude of gammas derived from official statistics with only slight variation in those from self-report studies. And, as Table 3 shows, the mean degree of association revealed by official data declines monotonically over the four decades from $-.73$ through $-.43$ and $-.22$ to a $+.04$ in the seventies, while from the fifties through the seventies it remains relatively constant at a low level for self-report instances ($-.04, -.11, -.03$). Thus the remarkable historical downward trend revealed by the data as a whole is actually because of declines registered by instances where officially recorded data were used to measure criminality. Furthermore, this historical trend is apparently not the result of other characteristics of the studies that might have varied over the decades. For example, Table 2 shows that this general pattern of declining associations has occurred within categories of offense, although not monotonically in every single situation. Thus instances where property crimes were studied show

variations from $-.35$ in the fifties to $-.09$ to $-.01$ while violent crimes show change from $-.52$ in the sixties to $-.06$ in the seventies. Similarly, controlling for the other characteristics of the studies does not affect the pattern of historical variation.

It is possible that the historical trend is an artifact of the volume and range of data that have been gathered in each decade. There are only three gammas available for the pre-1950 period, only ten in the fifties, but 170 in the sixties and 180 thus far in the seventies. It could be that the earlier data revealing high negative associations were simply unreliable and that as a greater range and number of data have been gathered the true low magnitude of the association has emerged. After all, the total number of instances by decade bears a monotonic negative relationship with the magnitude of the mean gamma, and the variance around the mean gamma has become smaller from decade to decade. But it is our opinion that changing reliability does not account for the historical pattern. This opinion is based on observations of the pattern of results. There is a consistent decline from decade to decade despite the magnitude of differences in number of instances on which the mean is based. For example, from pre-1950 to 1950 there is a substantial decline in mean gamma although the number of instances increases only from three to ten. And the mean gamma declines consistently from the sixties decade to the seventies although there is only a slight increase in the number of instances on which the figures are based (170 to 180), and that decline is about the same as between the fifties and sixties where there is an enormous increase in the number of instances used to calculate the mean (10 to 170). Furthermore, the trend is evident even when the number of cases declines from 46 official data instances in the sixties to only five in the seventies. Therefore we believe that the historical pattern of a declining association between social class and crime/delinquency is not simply a matter of reliability or of type of data.

A further possibility is that measures of social class and criminality have become more independent over time. Since earlier

data were produced by sociologists investigating social stratification, it is possible that an artificial negative relationship was inadvertently created. For example, if status is determined by community reputation or rating, then one's status may reflect noninvolvement with the police. And to the extent that reputational methods for determining class position declined historically as criminologists using "objective" criteria became more involved in the question, then the apparent class/criminality relationship would disappear as well. But this type of artifact could not account for the overall historical trend. The only instances included here where class was measured by rating methods involve data gathered before 1950. There was no tapering off of the use of rating methods over the four decades, hence this type of contamination might explain the initial change from before 1950 to the fifties but it could not account for continuation of the initial trend during the three most recent decades.

Fourth, it does not appear that this historical trend is the result of actual changes in the behavior of the various social classes. If it were, it would seem that measures based on self-reports would manifest temporal patterning. But as Table 3 and the regression analysis shows, the mean gamma for self-report instances has actually remained fairly constant from the fifties through the current decade. Since the historical decline is not a result of greater use of self-report data or of characteristics of the various studies, and is probably not the result of greater reliability or independence of the data or of actual changes in the behavior of the various classes, we are forced to conclude that the historical declines in association between social class and criminality must be because criminal justice agencies have changed the way they deal with members of the various classes. Indeed, this interpretation is supported by impressionistic consideration of historical events. It is reasonable to assume that the increasing consciousness of and publicized concern with individual civil rights, culminating with significant governmental action and landmark court decisions in the 1960s, has had some effect on the way criminal jus-

tice agencies conduct and report their business (see DeFleur, 1975). If publicity did make the police and lower courts more self-conscious of the rights of deprived members of the population, it would show up in official statistics. In the past, police no doubt could get away with closer surveillance of lower status persons and with making arrests on flimsier evidence than they now can. Moreover, expanded guarantees of due process surely have led to greater concern with judicial equality for all statuses. Certainly our empirical documentation of an accelerated decline in the class/criminality relationship in official statistics studies after 1964 lends plausibility to this interpretation.

But acceptance of this argument mandates an even stronger rejection of the class/criminality relationship hypothesis than that derived from the overall association of $-.09$ for the 363 research instances in the literature. Interpreting the historical trend as a result of changes in criminal justice biases implies, first, that contemporary instances based on official statistics more clearly reflect the true class/criminality relationship than did studies conducted in the past; that is, essentially no relationship (mean gamma = $+.04$). Second, if the relatively constant level of relationship revealed by self-report data is accepted as also being a fairly accurate reflector of the same class/criminality relationship (mean gamma = $-.06$), we must assume that the true relationship is somewhere between $+.04$ and $-.06$, or essentially zero. Third, this interpretation implies that the true relationship has remained consistently near zero and has only appeared to be greater because official data reflected biases in the law enforcement process which have now been ameliorated. In short, class and criminality are not now, and probably never were related, at least not during the recent past.

But this interpretation may not be accepted by all. Some will question whether self-report data are accurate enough to serve as a basis for establishing the true behavior of the classes over the three decades for which such data are available. Without this assumption, one would have to interpret the changes in the results of

official data studies as indicating actual changes in the behavior of the classes. Thus some will be willing to conclude only that social class is apparently not now related to crime/delinquency, although it may have been so related in the past. If this conclusion is correct, then it means that over the past three decades social class differences in life style, values, or constraints have become less important as predictors of criminality. Perhaps those of the higher classes have become less obedient to the law or those of the lower statuses have become more law abiding, or perhaps some of both changes have taken place. But regardless of the specific changes that may have accounted for this putative reduction in class differentiation, the net effect is consistent with findings by students of stratification. A number of scholars have commented on the decreasing significance of class in modern times (Cuber and Kenkel, 1954; Giddens, 1973; Jackson and Curtis, 1968; Landecker, 1960; Pfautz, 1953; Stone and Form, 1953). It is said that as advanced industrial societies become more culturally homogenized, the impact of stratum lessens. In a society where classification sorts people at birth into subcultures differing significantly in their definitions of reality, in their values, and in their socialization practices, class is likely to be a significant variable. But in a "massified" society where all but the extremes have the opportunity to accept or reject a single mass culture, class is likely to have significantly less impact. There are indications that industrial societies inexorably move, in Dennis Wrong's phrase, toward "inequality without stratification."

However, either of these interpretations of the data has serious implications for theories of deviance. If we interpret the findings to suggest that class differences in criminality actually have diminished, then we must acknowledge that theories of deviance have been time and culture bound. On the other hand, if the first conclusion—that class is not now and has not been related to criminality in the recent past—is accurate, then we have to question theories of deviance on more fundamental grounds. They are not time and culture bound but are simply erroneous. One may

not want to go that far, but it is difficult to avoid rather harsh imperatives. Certainly the general lack of predictive success of theories of deviance is consistent with this conclusion. Perhaps this failing is precisely because most deviance theories do rest on an assumption of class differences which has not in fact existed in recent history (Tittle and Villemez, 1977).

But whichever specific interpretation of the historical trend we choose, it is clear that contemporary data do not support the conclusion of a negative relationship between social class and crime/delinquency. Therefore it would seem that some shift in theoretical focus is called for. But that implication may be premature. The overall results of accumulated research, even of contemporary work, actually may be misleading. All studies included in our analysis were treated equally even though some were methodologically stronger than others. Although we controlled for some of these variations and found the results to be essentially unchanged, there is still the possibility that a negative relationship between class and criminality reflected in some of the studies is the "true" one, but is diluted by the inclusion of a number of other, perhaps less methodologically sophisticated pieces of research.

Second, none of the studies examined have adequately measured all the dimensions on which the classes might differ. Those of different status levels might vary in the cumulative probability of having done criminal or delinquent things by particular ages although not differing in the probability of committing misdeeds within a specified time interval (Gordon, 1976). They might differ in the ages at which criminal acts are first done or in the length of the period in a life cycle in which criminal or delinquent acts occur. Or they could differ in the total number of deviant acts committed by particular ages (Reiss, 1976). Furthermore, none of the studies takes into account the possibility of subtle variations in the ways in which deviant acts might be committed, nor do they include a broad enough range of kinds of offenses to capture all of the logically possible class variations. All of these might be especially important in accounting for the observed historical trends.

After all, if the character of social classes has been changing as many believe, the nature of current class influences on conformity might well be apparent only with precise and more focused measurements.

Nevertheless, it seems to us that the sociological community would do well to develop theories that emphasize variables which operate independently of supposed class differences. Such theories can be produced either by reconceptualizing the ones we have or by looking to new horizons. Actually most of the current class-rooted theories need not be so constrained. For example, members of all classes can experience inconsistency between goals and means which is contingent upon aspirations rather than objective circumstances. The key variable is not class position but rather discrepancy between aspiration and accomplishment, a variable which can vary nonsystematically over the class spectrum (see Keller and Zavalloni, 1964). Moreover, definitions favorable toward deviance are now widely distributed over the classes. Differential association may now be more a matter of the absence of countervailing interpersonal influences than of class position or place of residence. Similarly, other theories can be interpreted to emphasize generic processes rather than class related variables. Our findings suggest that this type of conceptualization is desirable.

But that may not be enough. We also need to identify more generic processes. What these processes might be we do not know, but we are confident that they will not be found as long as sociologists cling to the belief that almost everything ultimately can be reduced to a class variable.

SUMMARY

Studies examining the relationship between social status and crime/delinquency were reduced to comparable statistics using instances where the relationship was studied for specific categories of age, sex, race, place of residence, data type, or offense as units of analysis. The overall mean association (gamma) for 363 instances was found to be only $-.09$. In addition, the magnitude of association was

found to vary by type of data examined in the study as well as the decade in which the study was conducted.

Accumulated data suggest that for the past four decades there has been a monotonic decline in association between social class and crime/delinquency, with contemporary (those done since 1970) self-report and official statistics studies finding essentially no relationship between class and crime/delinquency. Moreover, these historical changes are found to be attributable to changes in findings by studies using official data. Further, analysis reveals a pattern of results which can be interpreted in either of two ways. One interpretation, contingent on confidence in the validity of self-report data, is that data never did demonstrate a negative relationship between status and crime/delinquency, and that in previous decades research appeared to show such a relationship because of biases in the criminal justice process which now have been corrected. Another interpretation, contingent upon confidence in the validity of official data, is that a class relationship did exist in the past, but no longer exists because social class generally has become less important.

But whichever interpretation is accepted, the implications undermine the purported class/criminality relationship which has fueled so much theoretical activity in sociology. Thus, numerous theories developed on the assumption of class differences appear to be based on false premises. It is now time, therefore, to shift away from class-based theories to those emphasizing more generic processes.

REFERENCES

- Akers, Ronald L.
1964 "Socio-economic status and delinquent behavior: a retest." *Journal of Research in Crime and Delinquency* 1:38-46.
- Arnold, William R.
1965 "Continuities in research: scaling delinquent behavior." *Social Problems* 13:59-66.
- Belson, William A.
1975 *Juvenile Theft: The Causal Factors*. London: Harper and Row.
- Berger, Alan S. and William Simon
1974 "Black families and the Moynihan report: a research evaluation." *Social Problems* 22:145-61.
- Bonger, William Adrian
1916 *Criminality and Economic Conditions*. Boston: Little, Brown.
- Box, Steven and Julianne Ford
1971 "The facts don't fit: on the relationship between social class and criminal behavior." *Sociological Review* 19:31-52.
- Bytheway, W. R. and D. R. May
1971 "On fitting the 'facts' of social class and criminal behavior: a rejoinder to Box and Ford." *Sociological Review* 19:585-607.
- Cameron, Mary Owen
1964 *The Booster and the Snitch*. Glencoe: Free Press.
- Casparis, John and Edmund W. Vaz
1974 "Social class and self-reported delinquent acts among Swiss boys." *International Journal of Comparative Sociology* 14:47-58.
- Chambliss, William J.
1975 *Criminal Law in Action*. Santa Barbara: Hamilton.
- Chilton, Roland J.
1964 "Continuity in delinquency area research: a comparison of studies for Baltimore, Detroit, and Indianapolis." *American Sociological Review* 29:71-83.
- Christie, Nils, Johs. Andenaes, and Sigurd Skirbekk
1965 "A study of self-reported crime." Pp. 86-118 in *Scandinavian Studies in Criminology*, Vol. 1. London: Tavistock.
- Clark, John P. and Eugene P. Wenninger
1962 "Socio-economic class and area as correlates of illegal behavior among juveniles." *American Sociological Review* 27:826-34.
- Cohen, Albert K.
1955 *Delinquent Boys: The Culture of the Gang*. New York: Free Press.
- Cohen, Albert K. and James F. Short, Jr.
1971 "Crime and juvenile delinquency." Pp. 89-146 in Robert K. Merton and Robert Nisbet (eds.), *Contemporary Social Problems*. Chicago: Harcourt, Brace, Jovanovich.
- Cohen, Bernard
1969 "The delinquency of gangs and spontaneous groups." Pp. 61-111 in Marvin Wolfgang and Thorsten Sellin (eds.), *Delinquency: Selected Studies*. New York: Wiley.
- Conger, John Janeway and Wilbur C. Miller
1966 *Personality, Social Class, and Delinquency*. New York: Wiley.
- Cressey, Donald R.
1966 "Crime." Pp. 136-92 in Robert K. Merton and Robert A. Nisbet (eds.), *Contemporary Social Problems*. Chicago: Harcourt, Brace.
- Cuber, John and William Kenkel
1954 *Social Stratification in the United States*. New York: Appleton-Century-Crofts.
- Cutright, Phillips
1968 "Reply to Glenn." *Social Forces* 46:539-40.

- DeFleur, Lois B.
1975 "Biasing influences on drug arrest records: implications for deviance research." *American Sociological Review* 40:88-103.
- Dentler, Robert A. and Lawrence J. Monroe
1961 "Social correlates of early adolescent theft." *American Sociological Review* 26:733-43.
- Doleschal, Eugene
1970 "Hidden delinquency." *Crime and Delinquency Literature* 2:546-72.
- Douglas, J. W. B., J. M. Ross, W. A. Hammond, and D. G. Mulligan
1966 "Delinquency and social class." *British Journal of Criminology* 6:294-302.
- Elliott, Delbert S.
1962 "Delinquency and perceived opportunity." *Sociological Inquiry* 32:216-27.
- Empey, Lamar T.
1967 "Delinquency theory and recent research." *Journal of Research in Crime and Delinquency* 4:28-42.
- Empey, Lamar T. and Maynard L. Erickson
1966 "Hidden delinquency and social status." *Social Forces* 44:546-54.
- Empey, Lamar T. and Steven G. Lubeck
1971 *Explaining Delinquency*. Lexington: Heath Lexington.
- Erickson, Maynard
1973 "Group violations, socioeconomic status and official delinquency." *Social Forces* 52:41-52.
- Garrett, Marcia and James F. Short, Jr.
1975 "Social class and delinquency: predictions and outcomes of police-juvenile encounters." *Social Problems* 22:368-83.
- Giddens, Anthony
1973 *The Class Structure of the Advanced Societies*. New York: Harper and Row.
- Glenn, Norval D.
1967 "Massification versus differentiation: some trend data from national surveys." *Social Forces* 46:172-80.
- Glenn, Norval D., J. P. Alston, and David Weiner
1970 *Social Stratification: A Research Bibliography*. Berkeley: Glendessary.
- Gold, Martin
1966 "Undetected delinquent behavior." *Journal of Research in Crime and Delinquency* 3:27-46.
- Gordon, Robert A.
1967 "Issues in the ecological study of delinquency." *American Sociological Review* 30:927-44.
1976 "Prevalence: the rare datum in delinquency measurement and its implications for the theory of delinquency." Pp. 201-84 in M. W. Klein (ed.), *The Juvenile Justice System*. Beverly Hills: Sage.
- Gould, Leroy C.
1969 "Who defines delinquency: a comparison of self-reported and officially reported indices of delinquency for three racial groups." *Social Problems* 16:326-36.
- Gove, Walter R. (ed.)
1975 *The Labelling of Deviance: Evaluating a Perspective*. New York: Wiley.
- Green, Edward
1970 "Race, social status and criminal arrest." *American Sociological Review* 35:476-90.
- Hardt, Robert H. and Sandra J. Peterson
1968 "Neighborhood status and delinquency activity as indexed by police records and a self-report survey." *Criminological* 6:37-47.
- Havighurst, Robert J., Paul Hoover Bowman, Gordon P. Liddle, Charles V. Matthews, and James V. Pierce
1962 *Growing Up In River City*. New York: Wiley.
- Hewitt, John P.
1970 *Social Stratification and Deviant Behavior*. New York: Random House.
- Hirschi, Travis
1969 *Causes of Delinquency*. Berkeley: University of California Press.
1972 "Social class and crime." Pp. 503-20 in Gerald W. Thielbar and Saul D. Feldman (eds.), *Issues in Social Inequality*. Boston: Little, Brown.
- Hollingshead, August B.
1947 "Selected characteristics of classes in a middle western community." *American Sociological Review* 12:385-95.
- Hood, Roger and Richard Sparks
1970 *Key Issues in Criminology*. New York: McGraw-Hill.
- Jackson, Elton F. and Richard F. Curtis
1968 "Conceptualization and measurement in the study of social stratification." Pp. 113-49 in H. M. Blalock and A. B. Blalock (eds.), *Methodology in Social Research*. New York: McGraw-Hill.
- Johnstone, John W. C.
1977 "Social class, social areas and delinquency." Unpublished paper, Institute for Juvenile Research, Chicago.
- Keller, Suzanne and Marisa Zavalloni
1964 "Ambition and social class: a respecification." *Social Forces* 43:58-70.
- Kelly, Delos
1975 "Status origins, track position, and delinquent involvement: a self-report analysis." *Sociological Quarterly* 16:264-71.
- Kelly, Delos and William T. Pink
1975 "Status origins, youth rebellion, and delinquency: a reexamination of the class issue." *Journal of Youth and Adolescence* 4:339-47.
- Kvaraceus, William C.
1944 "Juvenile delinquency and social class." *Journal of Educational Sociology* 18:51-4.
- Landecker, Werner S.
1960 "Class boundaries." *American Sociological Review* 25:868-77.
- Lander, Bernard
1954 *Towards an Understanding of Juvenile Delinquency*. New York: Columbia University Press.
- Laumann, Edward O.
1966 *Prestige and Association in an Urban Community*. Indianapolis: Bobbs-Merrill.

- McDonald, Lynn
1969 *Social Class and Delinquency*. Hamden: Archon.
- Masters, Stanley H.
1975 *Black-White Income Differentials: Empirical Studies and Policy Implications*. New York: Academic.
- Meade, Anthony
1973 "Seriousness of delinquency, the adjudicative decision and recidivism—a longitudinal and configuration analysis." *Journal of Criminal Law and Criminology* 64:478–85.
- Merton, Robert K.
1968 *Social Theory and Social Structure*. Glencoe: Free Press.
- Miller, Walter B.
1958 "Lower class culture as a generating milieu of gang delinquency." *Journal of Social Issues* 14:5–19.
- Nisbet, Robert A.
1959 "The decline and fall of social class." *Pacific Sociological Review* 2:11–7.
- Nye, F. Ivan, James F. Short, Jr., and Virgil J. Olson
1958 "Socioeconomic status and delinquent behavior." *American Journal of Sociology* 63: 381–9.
- Ossowski, Stanislaw
1963 *Class Structure in the Social Consciousness*. New York: Free Press.
- Pfautz, Harold W.
1953 "The current literature on social stratification: critique and bibliography." *American Journal of Sociology* 58:391–418.
- Polk, Kenneth, Dean Frease, and F. Lynn Richmond
1974 "Social class, school experience, and delinquency." *Criminology* 12:84–96.
- Reid, Sue Titus
1976 *Crime and Criminology*. Hinsdale: Dryden.
- Reiss, Albert J., Jr.
1976 "Settling the frontiers of a pioneer in American criminology: Henry McKay." Pp. 64–88 in J. F. Short, Jr. (ed.), *Delinquency, Crime, and Society*. Chicago: University of Chicago Press.
- Reiss, Albert J. and Albert Lewis Rhodes
1961 "The distribution of juvenile delinquency in the social class structure." *American Sociological Review* 26:720–32.
- Robinson, W. S.
1950 "Ecological correlations and the behavior of individuals." *American Sociological Review* 15:351–7.
- Rose, Arnold
1958 "The concept of class and American sociology." *Social Research* 25:53–69.
- Rossides, Daniel W.
1976 *The American Class System*. Boston: Houghton Mifflin.
- Shaw, Clifford R. and Henry D. McKay
1969 *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press.
- Slatin, Gerald T.
1969 "Ecological analysis of delinquency: aggregation effects." *American Sociological Review* 34:894–907.
- Stinchcombe, Arthur L.
1964 *Rebellion in a High School*. Chicago: Quadrangle.
- Stone, Gregory P. and William H. Form
1953 "Instabilities in status: the problem of hierarchy in the community study of status arrangements." *American Sociological Review* 18:149–62.
- Tittle, Charles R. and Wayne J. Villemez
1977 "Social class and criminality." *Social Forces* 56:474–502.
- Voss, Harwin L.
1966 "Socio-economic status and reported delinquent behavior." *Social Problems* 13:314–24.
- Walberg, Herbert J., Elaine Gee Yeh, and Stephanie Mooney Paton
1974 "Family background, ethnicity, and urban delinquency." *Journal of Research in Crime and Delinquency* 11:80–7.
- Warner, W. Lloyd and Paul S. Lunt
1941 *The Social Life of a Modern Community*. New Haven: Yale University Press.
- Wheeler, Stanton
1966 "Delinquency and crime." Pp. 201–67 in Howard S. Becker (ed.), *Social Problems: A Modern Approach*. New York: Wiley.
- Williams, Jay R. and Martin Gold
1972 "From delinquent behavior to official delinquency." *Social Problems* 20:209–29.
- Winslow, Robert W.
1967 "Anomie and its alternatives: a self-report study of delinquency." *Sociological Quarterly* 8:468–80.
- Wolf, Preben
1962 "Crime and social class in Denmark." *British Journal of Criminology* 13:5–17.
- Wolfgang, Marvin E., Robert M. Figlio, and Thorsten Sellin
1972 *Delinquency in a Birth Cohort*. Chicago: University of Chicago Press.
- Won, George and George Yamamoto
1968 "Social structure and deviant behavior: a study of shoplifting." *Sociology and Social Research* 53:44–55.

SEX DIFFERENCES IN SOVIET EARNINGS*

MICHAEL SWAFFORD

Vanderbilt University

American Sociological Review 1978, Vol. 43 (October):657-673

Although the Soviet regime is proud of its considerable efforts to bring women into the labor force on an equal footing with men, it publishes virtually no statistics which shed light directly on the extent to which sex differences in earnings still exist. This article examines some unpublished survey data which not only document the magnitude of these differences, but which also allow us to test some common hypotheses concerning factors giving rise to them.

For those investigating sex differences in occupational attainment, the Soviet Union affords an interesting case. For six decades Soviet policy has aimed at bringing women into the labor force on a large scale. Ideologically, this objective took root in Engels's observation that women would achieve social equality only when they ceased acting as servants of their families and obtained gainful employment on an equal footing with men. And it drew sustenance from Lenin's emphasis on expanding the opportunities available to women: "To involve the woman in socially productive labor, to free her from 'household slavery,' to liberate her from subjection—stupifying and humiliating—to the eternal and exclusive surroundings of kitchen and children: this is a major objective" (Sakharova, 1973:8, 14).

* Address all communications to: Michael Swafford; Department of Sociology/Anthropology; Vanderbilt University; Nashville, TN 37235.

I wish to express my appreciation to several agencies which have generously supported me with funds or facilities during the course of my research on this topic: the International Research and Exchanges Board, the Vanderbilt University Research Council, the American Council of Learned Societies and the Social Science Research Council (for a grant for Soviet Studies), and the Kennan Institute for Advanced Russian Studies. I also would like to express my gratitude to Dr. Murray Feshbach (of the Foreign Demographic Analysis Division of the U.S. Department of Commerce) for his assistance with sources, and to several scholars who offered useful comments on an earlier version of this paper: Professors Ivar Berg, Richard B. Dobson, Walter Gove, Alastair McAuley, and an anonymous referee. I would like to acknowledge my great debt to those Soviet scholars whose considerable labor yielded the data reanalyzed in this article.

Numerous measures consonant with this ideology have been adopted in the effort to facilitate women's professional advancement: child care centers have grown to accommodate over 30% of preschoolers, and schools with prolonged hours occupy some of those children whose employed parents cannot make other afternoon baby-sitting arrangements (Kotlyar and Turchaninova, 1975:113-4; Sakharova, 1973:63); women enjoy equal access to (and participation in) institutions of higher and special secondary education; they may return to their jobs after as much as a year's maternity leave; they are provided time off to nurse infants, and extra sick leave to care for ill children; and laws guarantee equal pay for equal work, as well as equal opportunity in other respects (Dodge, 1966:72-4; Rzhantsyna, 1971:16). These measures—as well as the difficulty of supporting a family on one income¹—have achieved a notable effect: women constitute more than 51% of the civilian work force, and though most

¹ The Soviet government designates families with less than 50 rubles' per capita monthly income as having less than enough to provide for the *minimum* standards of material well-being. A family of four, then, would require a minimum of 200 rubles during a typical month. But the average Soviet worker earned only 122 rubles in 1972 (TsSu, 1973a:516). (Since 1974 monthly welfare payments of twelve rubles per child under eight years of age have been given to families below the poverty line; see Kotlyar and Turchaninova, 1975:21).

Please bear in mind that (in keeping with Soviet practice) all financial statistics in this article exclude collective farmers and workers in subsidiary farming (who receive substantial payments in kind). They also exclude illegal income.

interrupt their employment at least briefly for childbearing, few make a career of being a housewife (Sakharova, 1973:68-79; Kotlyar and Turchaninova, 1975:16).

Western observers have exhibited a certain fascination with women's prominent role in the Soviet labor force, and feminists have found useful examples in Soviet efforts to enhance the career prospects of women. But our understanding of the effects of Soviet ideology and policy on the fares of women workers has remained shallow. True, it is generally known that more than 70% of Soviet physicians are women, and that women are found in other occupations (such as construction work) which are typically male domains in the West. Such scattered examples, however, cannot fully satisfy our curiosity about the quality of women's labor experience.

Two questions come to mind in this regard: (1) What systematic sex differences exist in the distribution of workers across sectors of the economy and over various levels of skill or responsibility? and (2) To what extent do women's earnings differ from those of men, and what factors account for these differences? Thanks to an abundance of Soviet census data on female employment, the first question already has received substantial attention (Dodge, 1966; 1977; Sacks, 1976; McAuley, 1978; Sakharova, 1973; Kotlyar and Turchaninova, 1975). We shall address it only insofar as it relates to the second question as a plausible explanation of differential earnings. Suffice it to say, there remain substantial differences between the distributions of men and women throughout the Soviet occupational structure, but the differences are generally less pronounced than those found in other countries.

The question of differential earnings, on the other hand, has attracted little investigation—primarily because data are scarce. A Soviet scholar has written that the sex of employees is ignored when statistics on qualifications, earnings, and training are gathered routinely by government agencies (Mikhailyuk, 1970:81). This claim has been corroborated by Western economists attempting to uncover separate statistics on men's and

women's earnings (Chapman, 1977:225). One authority, in discussing the dearth of Soviet statistics on income distribution, has concluded:

Soviet economists and Westerners alike must draw heroic conclusions from indirect and excessively aggregated data. . . . And nothing at all can be said as to whether equal pay for equal work prevails with respect to the earnings of males and females. One can merely observe that, as in market economies, there is a high inverse correlation between the percentage of women employed and the level of wages by sector of the economy. (Schroeder, 1972:313)

The data to be reanalyzed in this article, then, are unique: they were collected in a sample survey (in 1963), the results of which are recorded in voluminous cross tabulations in an unpublished Soviet dissertation; and they shed light directly on differential earnings.

Before turning to this evidence, however, we should review two tenets of Soviet wage policy, for doing so will bring the issues into sharp relief. The first principle is that wages should vary according to the quantity and quality of work performed. "From each according to his ability, to each according to his *work*" (i.e., amount and kind of work) is a common expression which embodies the spirit of current policy. According to a publication written to enlighten the working masses: "Distribution [of wages] according to work is an economic law of socialism. . . . This is what assures workers' interest in the fullest, most efficient use of work time" (Batkaev, 1972:6).

A corollary of this principle is that there should be equal pay for equal work. Soviets are prone to claim that their system is distinguished from Western systems in this regard. "Our country is the first to put into practice the principle of equal pay for equal work by men and women," reports Karapetyan (1968a:115). "For the socialist system the implementation of human rights is inseparable from the equality of women in every sphere of society's life and activity. In many Western countries, there is still discrimination in women's salaries" (*Pravda*, 1977; see also Kharchev and Golod, 1971:16-7).

The second principle is that wages

should be highest in those sectors of the economy whose further development is most urgent. "Societal demand—this is a completely new economic concept. . . . The system of remuneration should be constructed in such a way as to materially stimulate the distribution and utilization of labor resources to maximize outcomes for society" (Batkaev, 1972:15). The second principle, then, introduces a market mechanism (though Soviet ideologists would wish to avoid such terminology and others would question whether it functions effectively). It serves to justify (or perhaps rationalize) a substantial variation in wages across Soviet industries.²

Both principles come to play in discussions of sex differences in Soviet earnings. Soviet social scientists concede that women earn less than men on the average—though they offer little evidence on the magnitude of the difference. Sometimes, they attribute the disparity to women's lower qualifications and productivity in factory work (Kotlyar and Turchaninova, 1975:12, 57–84). Often, however, they explain that women's work is lighter and less intense:

As is well known, there exist many occupations which require primarily men's or women's work. As examples, we need merely point to such occupations as mechanic, miner, steel founder, driver, and others which are almost all in the sphere of men's activity. The work of these and similar occupations is complex, heavy, demanding and at times requires great physical strength; that gives rise to higher pay. On the other hand, the occupations of weaver, typist, and library worker as a rule are filled by women. Although these occupations demand significant concentration and attention, all the same their work is comparatively light, which explains women's relatively low level of pay. (Karapetyan, 1968a:116–7; see also Sakharova, 1973:32)

² A celebrated example of its use is the transportation industry's effort to alleviate its chronic shortage of bus drivers (generated in part by the fact that only a small percent Soviets have occasion to learn to drive). Though official wage scales are closed to public inspection—except insofar as people find them out when applying for work, or when discussing their wages with one another—this position is vigorously advertised for up to 280 rubles a month, i.e., substantially more than a physician's official salary after 15 years' service.

Thus, women tend not to work in the heavy industries, in which wages have long been relatively high. The negative correlation between *mean earnings* and *percent employed in a sector who are women* is easily seen in Table 1. In transportation, where women constituted only 24% of those employed in 1970, monthly earnings averaged 137 rubles; in health services, where 85% were women, monthly earnings amounted to only 92 rubles. Data aggregated in this fashion do not, of course, necessarily prove that women earn less than men. Nevertheless, heretofore such figures have served as the primary indication of this fact (Chapman, 1977; Schroeder, 1972:313). Thus, for lack of data even the most rudimentary questions have not been answered well. How great is the difference in earnings? To what extent *can* the difference be explained by women's employment in industries offering low pay, as Soviet economists have suggested? Do men and women with equal training earn equal salaries in those sectors of the economy where physical strength is not required?

Table 1. Mean Wages and Percent Female in 12 Soviet Economic Sectors: 1970

	(1) % employed who are fe- male: 1970	(2) Mean monthly wages: (rubles)
Transportation	24%	136.70
Construction	29	149.90
Agriculture (excludes collective farmers)	45	101.00
Science and scientific services	47	136.80
Industry	48	133.30
Housing and personal services; communal economy	51	94.50
Government administration	61	122.20
Communications	68	96.80
Education and culture	72	105.80
Trade, public dining, sales and procurement, material-tech- nical supply	75	95.10
Credit and insurance organizations	78	111.40
Health services	85	92.00
All sectors	51	122.00

Source: col. 1: Sakharova (1973:16); col. 2: TsSU, (1973a: 516–7).

Note: the (ecological) correlation between entries in cols. 1 and 2 is $-.73$.

The following analysis will shed light on these and other such issues.

STATISTICAL ANALYSIS

The evidence analyzed here is the by-product of a methodological study on estimating family income distributions conducted in 1963 by the Laboratory of Economic-Statistical Research of the Academy of Sciences of Soviet Armenia. The survey took place in Yerevan, capital of the republic, and one of about three dozen Soviet cities with more than a half million inhabitants. The question of how well these results represent the rest of the USSR, as well as of what changes might have transpired since the study, will be taken up later. One should bear in mind, however, that since a single set of job descriptions and wage scales for the entire country is maintained by the State Committee for Labor and Wage Problems, the wage structure in Yerevan is in principle the same as that of the rest of the USSR. (In some regions all wages are multiplied by a regional coefficient to compensate for such factors as a higher cost of living.) Moreover, Yerevan has for decades been subject to the same labor laws and regulations as govern the rest of the country.

A sample of 3,600 families was drawn, in which 4,841 adults reported independent sources of money income during the previous month (from wages, bonuses, student stipends, pensions, and the like). Each of these individuals was further described by five attributes deemed likely to serve as efficacious predictors of earnings: sex, age, formal education, level of skill or responsibility required by the individual's occupation, and sector of the economy in which he or she worked. Respondents were then cross-classified according to these five attributes, and a conditional mean income was calculated for those in each cell of the resulting five-way table—provided that at least three workers were classified in the cell.³

The dissertation in which these tables appear (Karapetyan, 1968a) includes only workers between age 25 and retirement (55 for women, 60 for men); age is dichotomized at 40. Education is trichotomized in a way that is reasonable in the Soviet context: completed less than seven years of general education; seven to ten years; and special secondary or higher education. Five levels of skill or responsibility are defined, ranging from unskilled labor to management. These are described fully in the notes to Table 2, where the breakdown of industries also is given. Excluding students and pensioners, 3,320 workers between the ages of 25 and retirement were included in the original sample. Of these, 146 (4.4%) are missing from the following analysis for one of two reasons: Karapetyan did not record cells containing less than three workers; and two cells were thrown out because he recorded markedly inconsistent entries for them in different parts of the dissertation. The data, then, consist of mean earnings in 227 cells of the five-way table, calculated on the basis of 3,174 gainfully employed people.

It is instructive to begin by examining some statistics characterizing the dependent variable (Table 2). Mean earnings, conditional on each of the four predictor variables other than sex, manifest the expected trends (col. 2). Those older than 40 earned somewhat more than those younger. As regards education, those with specialized training earned upwards of 40% more than those with less formal education. Mean earnings by industry followed a pattern similar to that in standard

seems likely that unskilled and poorly educated workers were undersampled. Also, individuals rather than families should have been sampled for that part of the study designed to model individual earnings. These shortcomings, however, present no insurmountable problem in this context—assuming that those included in the sample are not systematically different from the population of those with the same combination of socioeconomic characteristics. The unstandardized regression coefficients used in the following analysis will be little affected by any errors in the variance of variables introduced by unrepresentative sampling. Tests of significance (and standard errors) will, however, be somewhat affected by these errors—but they play no important substantive role in this analysis.

³ Karapetyan (1968b:19) reports that the sample was found to be representative when measured against census statistics for Yerevan. Unfortunately, detailed data with which to verify this claim are not included in the published census series. However, it

statistical handbooks of appropriate vintage, though direct comparisons are complicated by idiosyncracies in the categorization of transportation and industry in this study (TsSU, 1970). Finally, positions requiring skill and responsibility unsurprisingly brought the highest incomes, with managers averaging about twice the earnings of their most lowly subalterns.⁴

The statistic of greatest interest, however, is the 40.3-ruble disparity between the earnings of men and women—the latter earning only 65% as much as the former. This gap takes on the dimensions of a chasm when one considers that only 5% of the women earned as much as the average man.

Fortunately, the data allow us to test the most common explanations for this disparity: that women hold positions of less skill and responsibility than men on the average, and that they tend to work in poor industries. We can also take into account differences in the age structure of workers of the two sexes. Women in Yerevan's labor force are younger (and thus have less work experience) than men on the average, since older women (with larger families and more traditional values) are less likely to work than older men. And we can allow for the *greater* educational attainment of women in Yerevan's work force (TsSU, 1972b:517).

One approach to controlling for these factors would be merely to compare the earnings of men and women matched (simultaneously) on age, education, industry, and level. A less cumbersome and more informative technique is to regress the conditional mean earnings in each cell on the values of the five predictor variables characterizing that cell. The coefficients of such a regression model (which explains 81% of the variance of mean incomes across the 227 cells) are reported in Table 2 (cols. 4 and 5).⁵ They straightway

yield an answer to the question of whether the observed sex difference in earnings can be attributed to other factors included in the equation: for the most part, they cannot. The algebraic difference between the coefficients for men and women, -30.6 rubles, is an estimate of the direct effect on earnings of being female—and it amounts to three-quarters of the 40.3-ruble difference in conditional means observed above.

This estimate of the direct effect of sex is essentially an average over eight sectors of the Soviet economy. Such an average might well obscure some enlightening differences between industries. Soviet rationales for differential earnings, after all, sometimes attribute women's lesser earnings to their inability to perform heavy work as well as men. This implies that sex differences in earnings should be least in those sectors of the economy where heavy physical labor is uncommon, i.e., in science, government, education, and health care. This hypothesis was tested by adding sex-industry interaction terms to the regression model already discussed. These terms were found to provide a small but statistically significant increment (1.5%) in variance explained. In particular, the direct sex effect among employees in science and also in textiles, light industry and foodstuffs were found to differ substantially from the 30.6-ruble average. However, the direction of sex-industry interactions was not consistent with the hypothesis.

A convenient way of observing this is to compare regression coefficients in equations estimated separately for men and women (Table 2, cols. 6 and 7). Notice first that the disparity between regression constants for the two equations is virtually identical to the direct sex effect observed above (as it should be). More to the point, notice also that the coefficient for science is much larger for men than for women (27.83 vs. 5.99). Adding this difference to the difference in regression constants just pointed out reveals that in science the direct effect of being female is to put one at

⁴ Those interested in how training, years of service, industry, and level of skill or responsibility bear on wages should avail themselves of several monographs based on rare data, written under the auspices of the Foreign Demographic Analysis Division of the Department of Commerce: Rapawy, 1970; Hoffberg, 1968 and 1969; and Cerniglia, 1967.

⁵ A somewhat more efficient model is obtained by using the logarithms of earnings. Such a model ex-

plains 86% of the variance of the logarithms of the mean incomes in the 227 cells. The coefficients of this model are recorded in Table 2, cols. 8 and 9.

Table 2. Individual Earnings: Yerevan, 1963

	Univariate statistics			Linear* model		Linear model for each sex [†]		Log-linear* model	
	(1) N	(2) Mean earnings	(3) s.d.	(4) Regression coef.	(5) St. error	(6) Men's reg. coef.	(7) Women's reg. coef.	(8) Regression coef.	(9) St. error
All respondents	3174	98.28	36.16	93.83	—	108.21	77.89	4.45	—
I. Sex									
Regression constant									
Female	1284	74.29	23.04	-15.30	1.24			-.167	.010
Male	1890	114.59	34.33	15.30	1.24			.167	.010
II. Age									
25-39	1974	93.50	29.12	-5.03	1.17	-6.34*	-2.81	-.035	.010
40-retirement ^a	1200	106.15	44.33	5.03	1.17	6.34	2.81	.035	.010
III. Education ^b									
High	1554	116.88	38.14	12.07	2.36	14.76*	6.84	.122	.020
Medium	1256	82.86	22.40	-2.69	1.86	-4.57	-.59	-.016	.016
Low	364	72.11	21.90	-9.38	n.a.	-10.19	-6.25	-.106	n.a.
IV. Industry ^c									
1. Science	287	134.84	52.24	17.99	3.61	27.83*	5.99	.096	.031
2. Construction	165	112.25	23.10	3.19	4.43	1.90	7.13	.058	.037
3. Government administration	407	110.58	41.15	-.78	2.98	.64	-3.50	-.009	.025
4. Mining, metallurgy, chemical, rails, etc.	495	109.68	27.64	11.21	2.76	12.17	7.39	.131	.023
5. Machine & instrument building	400	99.76	25.22	+4.55	n.a.	2.55	6.16	.066	n.a.
6. Education and health care	583	84.46	29.61	-16.50	2.82	-14.05	-16.85	-.179	.024
7. Textiles, light industry, foodstuffs	372	79.23	21.15	-5.85	3.18	-14.74*	2.28	-.031	.027
8. Retail trade, communications, city transportation	465	79.18	26.58	-13.81	3.03	-16.30*	-8.60	-.132	.026
V. Level ^d									
1. Managers	432	142.32	39.49	28.73	2.99	29.57	24.88	.261	.025
2. Highly skilled specialists	829	116.83	30.36	14.45	n.a.	12.18	19.89	.206	n.a.
3. Moderately skilled employees; skilled production workers	1242	88.46	20.46	-8.43	2.05	-5.20	-8.63	-.050	.017
4. Unskilled production workers	354	64.10	21.95	-17.02	3.33	-20.89	-18.96	-.223	.028
5. Unskilled service workers	317	66.46	12.15	-17.73	3.43	-15.66	-16.78	-.194	.029

Source: Calculated on the basis of data contained in Karapetyan (1968a).

^a Provided they have worked 20 years, women may retire at 55; men at 60.^b Education is trichotomized as follows: higher and special secondary education; general secondary and seven year education (*semileitnoye*); elementary (*nachal'noye*) and without education. Karapetyan doesn't specify how he classified those who failed to complete a given level of education—at that level or at the next lower level. Judging from the distribution, which is high compared with census data (18SU, 1972b: 517), those who completed at least part of a

higher or special secondary education were categorized as high, and perhaps those who had completed grade five or six were categorized *medium* even though they had not completed "seven year education."

* A translation of Karapetyan's description of industries follows: (1) science; higher education; planning, investigative, and research organizations; (2) construction; (3) *upravleniye* (government administration, and probably includes credit and insurance organizations); (4) mining, metallurgy, electrochemical, chemical, construction materials, railroad transportation; (5) machine-tool, machine equipment, and instrument manufacturing; (6) lower education, health care, social insurance, culture, and sanitation; (7) textiles, light industry, and foodstuffs; (8) communications, municipal and automobile transportation, communal economy, procurement, wholesale and retail trade, and public dining.

* A translation of the definition of levels follows:

- (1) managers of institutions, enterprises, organizations, shops, sections, laboratories, sectors, divisions, and shifts;
- (2) specialists of high qualification;
- (3) specialists with middle qualifications; work superintendents; foremen; brigade leaders; technicians; skilled craftsmen;
- (4) unskilled workers, apprentices, minor service personnel (MOP), and miscellaneous workers [in production];
- (5) [unskilled service] workers in retail trade and public dining; office workers; service workers in places of entertainment; other service personnel.

* The mean income in each of the 227 cells was regressed on all five of the attributes characterizing the cell; the means were weighted by the number of workers on which they were based. The regression coefficients obtained are *identical* with those which would have been obtained had we regressed the income of each of the 3,174 workers on the predictor variables. The model explains 81% of the between-cells variance. Using the mean incomes in the cells eliminated within-the-cell variance (which is usually part of the error variance) from the total variance to be explained. Had we been using individual data, the variance explained would have been lower. Effects coding was used to represent the categorical predictor variables; hence, the effects within each of the five predictor variables sum to zero. (See Kerlinger and Pedhazur, 1973.) As usual to prevent linear dependencies, one category of each variable was omitted during the regression; for that reason, standard errors are not available in some cases. See the text for comments on first-order interactions.

* Coefficients separated by an asterisk are significantly different at the .05 level.

* The natural logarithm of the mean income in each cell was regressed on all five of the attributes represented by the cell; the means were weighted by the number of workers on which they were based. This model explains 86% of the between-cells variance.

a disadvantage of more than 50 rubles, rather than 30.6. Thus, in one of the sectors expected to have a smaller-than-average direct sex effect, quite the opposite is observed. A list of sectors ordered according to the (decreasing) magnitude of direct sex effects follows: science, mining, government administration, education, machine building, construction, retail trade, and textiles. It is evident that education and government administration are also among the sectors exhibiting unexpected larger-than-average direct sex effects. Clearly, these findings are not consonant with the hypothesis that men earn more than women by virtue of performing more physically demanding labor.

The sex-specific equations just examined also reveal sex-age and sex-education interactions. Especially notable is the interaction term involving higher education: for men the return for higher or special secondary education is approximately twice that for women. Restated in terms of a sex effect, among employees with higher or special secondary educations, the direct sex effect is 38 rubles rather than 30.6.

This discussion of interaction terms should not, however, be permitted to distract one from the central finding of the analysis—that most of the disparity between men's and women's earnings cannot be attributed to women's less advantageous distributions across sectors of the economy and over levels of responsibility. In fact, even if the distributions of men and women were rendered identical on the four predictor variables other than sex (i.e., even if both were made equal percentwise to the pooled distribution described in Table 2, col. 1), women would earn only 72% as much as men.⁶

⁶ Karapetyan's primary aim was to present an estimate of the income distribution of Yerevan's families using data from the 1959 census and from the survey reanalyzed in this article. He nevertheless devoted substantial attention to estimating the effects of the five predictor variables (without availing himself of regression). By comparing cells in which men and women were matched with respect to age, education, sector and level, he calculated that the effect of being male was to add about 39% to one's income (1968a:149). This estimate is quite consistent with that implied by the model presented in Table 2.

DISCUSSION

The substantial earnings differential just documented calls for explanation. In undertaking to provide one, we may begin by dismissing some factors which might well figure heavily under other circumstances. First, part-time employment (which accounts for about a third of the disparity in men's and women's earnings in the United States; Council of Economic Advisors, 1973) is scarce in the USSR, and probably contributes very little to the direct sex effect observed above. Also, differences in the number of hours worked monthly probably explain little of the disparity. True, disproportionate numbers of women work in health and in education, where the relatively poor wages perhaps reflect shorter authorized work days (Chapman, 1977). But this cannot explain the substantial sex effect found *within* health and education, as well as other industries.

Except as noted below, we may also set aside a series of factors which typically impede the professional advancement of women, and thus depress their earnings: interruptions for childbearing (shorter in the USSR than elsewhere, as a consequence of provisions outlined above; see Sakharova, 1973:69); lack of the mobility required to maximize advancement, reflecting familial considerations; and traditional attitudes which dictate that women should not manage (probably most prominent in the Central Asian republics). Such factors no doubt explain in part why women are somewhat less likely than men to gain positions in the upper reaches of the occupational hierarchy. But they cannot account for the direct sex effect uncovered by the regression analysis, since

(i.e., the reciprocal of the .72 just given in the text is 1.39).

The lengthy quotation rationalizing sex differences in earnings presented earlier in the article is one of Karapetyan's explanations for his findings. In the published summary of his dissertation (which offers no statistics), he also conjectures that *within* levels men occupy the more highly-paid positions (1968b:20). He does not attempt to put his findings into perspective by referring to other Soviet studies; nor does he suggest that the results from Yerevan might be atypical. Unfortunately, since Soviet regulations prohibit photocopying dissertations, I do not have the text of Karapetyan's dissertation available for more extensive explanation.

that coefficient for all practical purposes compares the incomes of men and women who (for whatever reason, and in whatever quantities) work at the *same* level, and in the same industry, with the same age and education.

This conclusion, however, must be tempered. A close look at the description of levels in Table 2 reveals hierarchies within the defined levels. For instance, the top level includes managers of factory shifts as well as managers of large enterprises. By virtue of greater responsibility, the latter would be expected to earn more than the former—yet both were treated as having the same level in the above statistical analysis. Since women may have been less likely than men to occupy the upper ranks *within* the levels defined in this study, more adequate measurement of level might have yielded a smaller estimate of the direct sex effect. In other words, errors in the measurement of level may have been correlated with sex, and thus effects which should have been attributed to level were instead laid to the account of sex.

But the merit of this argument should not be overemphasized. Level is fairly well measured, as examining the definitions and the conditional means in Table 2 confirms. Yet when *level* is added to a regression equation containing only *sex* as a predictor of earnings, the absolute value of the estimated sex effect diminishes only 5.7 rubles (from 40.3). Now, there is no reason to suppose that the problematic *unmeasured* differences in the distributions of men and women *within* levels are any greater than the *measured* difference of their distributions *over* the five levels (recorded in Table 3). And there is therefore no reason to expect that refining the measurement of level could reduce the estimated sex effect by more than another few rubles—certainly far less than the estimated 30.6-ruble direct sex effect in question.

Beyond this shortcoming in the measurement of level, we may point to women's lower productivity as an unmeasured factor perhaps biasing the estimate of the sex effect. Evidence of differential productivity has been gathered primarily for manufacturing industries where pro-

Table 3. Distribution of Men and Women over Five Skill Levels

	Level ^a					Total	N
	1 (high)	2	3	4	5 (low)		
Women	5%	31	35	17	11	99% ^b	1,284
Men	19%	22	42	07	09	99%	1,890

^a Level is defined in fn. d of Table 2.

^b Rows sum to less than 100% because of rounding error.

Source: based on data in Karapetyan (1968a).

ductivity is easy to measure (Kotlyar and Turchaninova, 1975:12, 57-94). But there are grounds for speculating that women's productivity is somewhat lower than men's in other sectors of the economy as well. Blame for this is usually laid to women's "double burden" or "second workday" (Karchev and Golod, 1971:31). One of the most common exercises of Soviet survey research, in fact, is to gather time budgets confirming anew that employed wives bear the brunt of household chores (Sakharova, 1973:46-66). To hark back to Lenin's terminology, women have indeed been brought into "socially productive labor"—but they remain saddled with the "kitchen and children" as well. Adding considerably to their burden has been an underdeveloped service sector and a dearth of household appliances. About the time of the survey in Yerevan, for instance, only some 5% of Soviet families owned refrigerators, and approximately 15% owned washers (Shumilova, 1964:124-5). Soviet planners have been attempting to improve women's lot, however, and by 1972 the percent of Soviet families owning refrigerators and small wringer washers had grown to 38% and 57%, respectively (Sakharova, 1973:61).⁷

Two other related factors which an ideal study would include are *job tenure* and *overall labor force experience*. These variables have proven useful in explaining some of the earnings differential in Western countries (see Sawhill, 1973; Bibb and

Form, 1977; and Mincer and Polacheck, 1974). Under some circumstances, one might utilize *age* as a makeshift surrogate for total labor force experience, but as a dichotomous variable it would certainly function poorly in such a role in this analysis. We can only speculate that these variables would serve less potently in explanations of differential earnings in the USSR than they have in other countries. Soviet women are less likely than Western women to move in and out of the labor market, and are therefore less likely to lose the benefits of continuity in employment. Moreover, their right to return to former positions after as much as a year's maternity leave would also serve to minimize the effects of breaks in employment due to pregnancy. Thus, it would seem that sex differences in labor force experience are probably not as pronounced in the USSR as elsewhere, and that such differences are therefore less likely to explain any earnings differential.

Finally, we cannot pass by the specter of bald discrimination. On the basis of mere averages of men's and of women's earnings in Western countries, Soviet authors are uniformly pleased to settle for such an explanation of the well-known sex differences in Western earnings (Karchev and Golod, 1971:16-7; Sakharova, 1973: 8-9). By this standard, Soviet women suffer a fate like that of their Western counterparts. And according to these data, the most common Soviet explanations of the disparity appear unsatisfactory. We shall return to this point in the conclusion.

COMPARISONS ACROSS REGIONS AND TIME

It has already been pointed out that wage scales, as well as labor laws, are consistent throughout the USSR. It never-

⁷ Various Soviet agencies also have undertaken to relieve men's aversion to domestic chores. The author witnessed an episode in the Ministry of Public Health's effort—a movie short informing men that the exercise of carrying groceries (after having first shopped for them) would improve their health. Judging from levity in the audience reaction, the advice fell short of its mark.

theless remains plausible that results from Yerevan are unrepresentative. The roles of Armenian women in 1963 were doubtless more traditional than those of women in several of the European republics (Katz, 1975). Fertility in Soviet Armenia, for instance, was substantially higher than the national average (Feshbach and Rapawy, 1976:124). And in 1959 the mean size of urban families in Armenia was 4.5 members—one member larger than the average for all Soviet families in urban places (TsSU, 1974:234–7).

One consequence of these facts is that Yerevan's women have been less likely than other urban women to enter the labor force. In 1959, the ratio of *number of urban women working in the socialized economy to number of urban women between ages 16 and 54* was .47 in Armenia, but .65 in the USSR. In 1970, the corresponding figures were .71 and .86.⁸ And in 1966 women constituted only 40% of the labor force in Armenia, but 50% in the USSR (TsSU, 1968: 77).

However, this disparity in likelihoods of employment is immaterial here. The estimated direct sex effect is not a function of the percent of women employed, i.e., the analysis concerns only people *in* the labor force. Also irrelevant would be evidence showing that Armenian women are less likely than women of other republics to advance through the five levels of skill and responsibility defined in this study. Such a difference *would* render the disparity be-

tween men's and women's mean earnings greater in Armenia than elsewhere, but it would not of its own accord yield a larger direct sex effect in Armenia. To reiterate a crucial point, the very aim of the multivariate analysis was to estimate a sex effect in such a manner that differences in men's and women's distributions over age, education, sector, and level would not contaminate it. By the same token, differences in these distributions across republics would not affect a similarly calculated (unstandardized) direct sex effect.

But again, this argument must be tempered to allow for the measurement problem discussed in the previous section. If sex roles are more traditional in Armenia than average, it seems plausible that employed women there would find it unusually difficult to advance to the upper echelons *within* the defined levels. To the extent that this measurement problem inflates the estimate of the direct sex effect, we would expect a less inflated coefficient in regions where women advance *within* these levels more easily, i.e., a coefficient with a smaller absolute value than 30.6.

In this regard, it is instructive to consider some Soviet census data which allow us to compare Armenian women's success in gaining top positions with that of women over the entire USSR. Table 4 (compare cols. 3 and 5, or 4 and 6) demonstrates that although Armenian women made up smaller than average proportions of employees in law and in top management, they were in fact about as likely as other women to constitute significant proportions of scientific workers, middle-level managers, medical workers, and engineering-technical personnel. Also, notice the similarities between the proportions of Armenian and other women within each of the hierarchies of scientific, management, and medical employees. In sum, the effects of the assumed traditional sex roles in Armenia appear only modestly greater than elsewhere, and we would therefore expect little difference in estimated direct sex effects on this account.

There is also another approach to evaluating how representative these results from Yerevan are. From time to time, bits of information on differential

⁸ These statistics are imperfect indicators of the percent of eligible women who are gainfully employed. Some employed women are over 54, yet women over retirement age are not included in the denominator. Also, many females over 15 (included in the numerator) are in school full time, and thus are ineligible for employment. Finally, women employed in subsidiary (private) farming are excluded; Armenian women are doubtless more likely than average to engage in subsidiary farming (thus depressing the ratios for Armenia) (TsSU, 1973b:6–12; 1972a:14, 15, 66, 67). Better estimates of the percent of eligible women gainfully employed appear from time to time, but I am not aware of any which would allow direct comparison of Armenia with the USSR.

Since Armenian statistics constitute a subset of USSR statistics, comparisons of the two are somewhat tainted by a lack of independence. However, since the population of Armenia amounts to only 1% of the USSR population, this lack of independence is inconsequential.

Table 4. Percent Employed Who Are Women: Armenia and the USSR

	Armenia		Standardized figures for Armenia ^a		USSR	
	1959 (1)	1970 (2)	1959 (3)	1970 (4)	1959 (5)	1970 (6)
All employed people	42	46	48	50	48	50
I. Scientific workers ^b	30	36	34	39	36	38
Academician	07	09	05	06	07	09
Dotsent	17	20	18	18	17	20
Senior research associates	29	29	38	26	29	29
Junior research associates	51	48	56	48	51	48
II. Top management	17	19	19	21	26	32
III. Middle management	9	12	10	13	13	16
Heads of enterprises (<i>Nachal'niki</i>)	8	8	12	11	12	13
Heads of sections (on state farms)	22	21	32	28	20	24
	14	14	20	19	15	16
IV. Engineering-technical personnel	33	37	38	40	40	44
Engineers	27	36	33	43	32	40
V. Medical personnel	83	84	95	91	89	89
Head doctors	38	48	41	51	54	53
Doctors	72	69	77	73	79	74
Physicians assistants and midwives	67	67	72	71	84	83
Nurses	98	98	100	100	99	99
VI. Workers in jurisprudence	14	16	16	17	33	38

^a The question being raised concerns whether employed women in Armenia are as likely as employed women in the USSR to obtain top positions in the occupational hierarchy. It would be misleading to compare statistics from Armenia (cols. 1 and 2) with those of the USSR (cols. 5 and 6) without adjusting for the fact that the percent of the Armenian work force which is female is smaller than the corresponding percent of the USSR work force. (An analogous comparison which illustrates the point would be comparing death rates in California and Florida without adjusting for the disproportionate number of retired people in Florida.) In columns 3 and 4, the Armenian statistics have been adjusted so that any differences between Armenia and the USSR cannot be attributed merely to the smaller percent of the Armenian labor force which is female. For example, 36% of scientific workers in Armenia (1970) were female (col. 2). The adjusted figure (col. 4) is 39%. This was calculated by multiplying 36% times the ratio of percent of all employed people in the USSR who are female to the percent of all employed people in Armenia who are women, i.e., 50%/46%. This same procedure was followed for all the main entries, i.e., those preceded by roman numerals. The indented entries (under I, III, and V) were standardized within categories. Thus, in 1970 25% of senior research associates in Armenia were female. This figure was adjusted by multiplying it by the ratio of percent of USSR scientific workers who were female to percent of Armenian scientific workers who were female, i.e. 38%/36%.

^b For *scientific workers*, the years are 1960 and 1968, not 1959 and 1970. Most scientific workers are below the four levels listed, i.e., they do not bear titles. Sources: TsSU, 1970: 143; and TsSU, 1968: 248.

Source for statistics in cols. 1, 2, 5, and 6, under categories II through VI: TsSU, 1973b: Tables 18 and 31.

earnings in other republics are published. Though these statistics often admit of only labored comparisons with those derived from the above multivariate analysis of survey data, they nonetheless leave some useful impressions.

Consider, for example, the findings of a survey conducted in 1969 to 1972 of 22 manufacturing enterprises located

throughout Soviet Russia (employing about 11,000 people in the production of clothing, processed foods, and instruments). Kotlyar and Turchaninova (1975:4, 139) report that only 21% of the women in these enterprises brought home more than half their family's wage receipts, while more than 62% of the men did so. Similar conclusions were reached

in a 1968 study of 650 families in four Kiev plants (representing light industry, textiles, and food processing). Only 8 to 18% of the women employed in these plants earned more than their husbands; another 9 to 21% earned approximately as much; and 71 to 79% earned significantly less. In these same plants, interestingly, women's earnings averaged 83% to 90% of men's (Sakharova, 1973:29-31). These would appear to be some of the smallest Soviet earnings differentials to come to light. They cannot, however, be taken as representative of these industries in Kiev, since Sakharova offers too little information on the selection of the four plants. Nor can they be taken as representative of all sectors of the economy: as the data from Yerevan illustrate, sex differences in earnings are unusually small in precisely these industries. More to the point, modest experience suggests that surveys of specific jobs in a limited number of firms uncover smaller earnings differentials than surveys of the population of workers do. For instance, although U.S. Census figures demonstrate that full-time female employees earn only about two-thirds as much as males (after adjusting for females' shorter hours), "... Department of Labor surveys have found that the differential almost disappears when men's and women's earnings are compared within detailed job classifications and within the same establishment" (Council of Economic Advisors, 1973: 106). This is not to downplay the validity or utility of either type of study. Rather, it is to emphasize that any effort to determine whether the survey results from Yerevan are representative should rely on similar surveys of the population of workers in other regions of the country, if possible.

Two or three massive surveys of this sort have been conducted by the Central Statistical Administration (TsSU). The results of these surveys are for the most part closed even to Soviet social scientists, but from time to time a few statistical crumbs find their way into print. Consider, for example, some data given by Rimashevskaya, whose writing on Soviet income distribution has gained high regard.⁹ She de-

finies the *first worker* of a family as the one earning the most money; all others are designated *second workers*. She reveals that, according to the well-known 1958 survey of 240,000 Soviet families, first workers brought in 61% of the money incomes of families; second workers brought in 32%; transfer payments amounted to the remaining 7% (1965:29-30). These figures would seem to indicate that second workers earned only half as much as first workers. However, there were only 80% as many second workers as first workers (since some households contained only one worker). Adjusting for this, we find that the average second worker earned 66% as much as the average first worker [$.32/ (.61 \times .80)$]. This is not a simple comparison of men's and women's earnings, of course, since some first workers were no doubt women. But all available evidence indicates that the odds are against wives earning more than their husbands. As stated in the previous paragraph, for instance, even in those industries manifesting the least earnings differential, less than a fifth of the wives brought home more than their husbands. And in Yerevan, only 5% of the surveyed women earned as much as the average man. Hence, it seems likely that most first workers in the above survey were men—and 66% provides a rough estimate of how women's earnings compared with men's. As such, its similarity to the estimate from Yerevan (65%) is striking. (See also Korzhenevskiy., 1965:80, for another study leading to a similar conclusion; and McAuley, 1978:7a).

Pointing the same direction are three statements published between 1968 and 1974 by Soviet economists who seem privy to government data. All three were addressing the question of how much each of two workers in a family of four would have to earn to provide minimum financial security:

(a) As data from budget research of workers [*rabochie*] in industry show there is a definite and characteristic relationship be-

⁹ See especially Rabkina and Rimashevskaya

(1972). Several recent studies of income distribution make heavy use of this book. See Wiles (1974) and McAuley (1977).

tween the wages of the heads of families and other members. Thus, in the set of families with per capita incomes of 50 rubles per month, the wages of the second worker constitute about 40% of the total wages. From this it follows that in a budget providing for minimum material well-being, the wages of the second worker must be set (minimally) at a level of no less than 60 rubles per month, and the head of the family, at about 90 rubles. (Sarkisyan and Kuznetsova, 1968: 75-6)

(b) As a rule, the first worker's wages constitute 60%, and the second worker's 40%, of the general sum of wages; therefore, to maintain a budget of material well-being, the earnings of the primary workers should be 90-100 rubles, and of the second, 60-70 rubles per month. (Kunel'skiy, 1972:64)

(c) . . . [T]he amount of total wages of the two given people must be equal to about 160 rubles per month, out of which one worker receives, say, 100 rubles and the second, 60 rubles. (Oganyan, 1974:89)

These quotations exhibit a sturdy consensus on the relative sizes of earnings brought home by family heads and second workers. But more than this, they reveal the extent to which this relationship is considered normal—even a basis for planning!

It should be added, in closing, that there is reason to suppose that the earnings differential might have diminished over the past two decades (but not necessarily the direct sex effect). Women have likely been the primary beneficiaries of two important developments in Soviet wage policy: the 160% increase in legal minimum wages (from 27 to 70 rubles monthly) has outstripped the percent of increase in mean earnings; and raises in health and in education (with high concentrations of women) have exceeded (percentwise) those in most other industries. No evidence confirming a shrinking differential has come to light, however.¹⁰

¹⁰ Soviet women can take some consolation, however, in the fact that they probably recoup some of their losses through greater pension benefits: they are entitled to retire five years earlier than men, and their life expectancy (at birth) is ten years greater than that of men (Baldwin, 1973:14). It is most unlikely, though, that Soviet officials will take refuge in such an explanation for an earnings differential which they do not acknowledge in public.

CONCLUSION

The rare survey data analyzed here paint a picture which is altogether familiar. In the USSR, as elsewhere, women's earnings fall significantly below men's. In the republic capital surveyed, women brought home only 65% as much as men, and other evidence suggests that this estimate was applicable to the entire Soviet Union during the sixties. One is struck by the similarity of this figure to statistics for Western industrial countries (Table 5).

This similarity obtains despite several relevant differences between the USSR and these other industrial countries. Marxist-Leninist ideology has stressed the liberation of women from domestic servitude throughout the 60-year history of the Soviet Union. Second, circumstances as well as ideology have compelled the USSR to make the most of its women's labor capacity. World War II brought the deaths of millions of Soviet soldiers; women filled the resulting void in the labor force. Also, efforts to industrialize the country have depended in part on recruiting housewives into the labor force. Even today, Soviet officials complain about shortages of specialists and skilled workers, and claim that there is no unemployment problem. Yet a third difference is that the Soviet Union has adopted many of the progressive measures which according to common sense should facilitate women's professional advancement. Fourth, it boasts one of the most thoroughly socialized economies in the world. This fact, together with Soviet managers' preoccupation with filling output quotas rather than minimizing labor and production costs, would lead one to expect the absence of the conventional profit motive for exploiting women attributed to Western capitalists by "radical" economists (Gordon, 1972:78-9. Finally, the centralized wage controls implemented in the USSR might have been expected to ease considerably the task of reducing any earnings differential.

This effort to explain sex differences in Soviet earnings has nevertheless yielded results reminiscent of similar studies in the United States. Less than a quarter of

Table 5. Earnings Differentials in Western Industrial Countries

Country	Women's earnings as a percentage of men's	Notes ^a
Belgium	73%	Average hourly earnings of adult workers in industry; October, 1975
France	69	Average annual earnings of full-time labor in private and semiprivate industry; 1973
Great Britain	71	Median gross hourly earnings of full-time adult manual workers; 1976
	62	The same, for nonmanual employees
The Netherlands	78	Average gross hourly earnings of adult industrial workers; 1975
Sweden	69	Average monthly salaries of adult white-collar employees in mining and manufacturing; 1976
	87	Average hourly earnings of labor in mining and manufacturing; 1976
Switzerland	66	Gross hourly earnings of skilled workers in industry and trade; 1976
	70	The same, for semiskilled and unskilled workers
	65	Gross monthly earnings of white-collar employees in all sectors; 1976
United States	61 ^b	Median earnings of full-time wage and salary workers over 25; May, 1976
West Germany	74	Gross hourly earnings of skilled workers in industry; 1976
	77	The same, for semiskilled workers
	83	The same, for unskilled workers
	67	Gross monthly earnings of technical employees in industry, trade, and financial institutions; 1976

^a The terms *worker* and *labor* exclude white-collar employees.

^b When allowance is made for the longer hours which full-time men work, this figure jumps to 66%, judging from previous analyses (Fuchs, 1971).

Source: the annual statistical yearbook for each country, listed in the reference appendix.

the disparity could be explained by differences in male and female distributions across levels and sectors of the Soviet economy and by differences in human capital (as measured here). Similarly, Featherman and Hauser (1976:481) could explain only 15% of the U.S. earnings differential with such variables (see also Bibb and Form, 1977; and Mincer and Polacheck, 1974).

Given the secrecy which shrouds Soviet wages, the prospects for fully explaining the earnings differential seem dim. Nevertheless, it appears that two avenues of research will prove fruitful as the necessary data become available. The first concerns the effects of marital roles on the

careers of women. Western studies have revealed that the earnings of *unmarried* women compare much more favorably with those of men than the earnings of married women do (Fuchs, 1971:10; Council of Economic Advisors, 1973: 105). There is reason to expect that the same would be found in the USSR. Soviet wives still perform most domestic chores, and most do still "serve" their families (Gordon and Klopov, 1975; Szalai, 1972). This obligation is a drag on their energy and productivity at work, and no doubt decreases the probability that they will seek demanding positions. We may also conjecture that the dearth of reliable and acceptable birth-control paraphernalia in-

creases women's uncertainty in pursuing careers. (Abortions are more common than births in parts of the USSR.) Also, of course, many women prefer their wifely roles, and thus may be more prone than men to take employment merely to fulfill social obligations or to supplement family income (Kharchev and Golod, 1971). In fact, the fear of losing women's full-time labor is no doubt one of the reasons why the Soviet government permits scarcely any part-time employment.

The other avenue of research deserving further investigation concerns labor-market segmentation. In recent years, some social scientists have challenged anew the adequacy of orthodox economic explanations of racial and sex differences in earnings (for reviews, see Cain, 1976; Gordon, 1972). In doing so, they have highlighted the existence of a dual labor market, in which minorities and women tend to be confined to a secondary market which does not demand employment stability, and which does not offer security, good pay, or opportunities for advancement. It is worth observing that Soviet policies in some ways encourage occupational segregation. Soviet regulations require, for instance, that girls in fourth through eighth grades be given occupational training for domestic or service labor, while boys receive skills suitable for industrial labor. Regulations on blue-collar job training specify the sex of appropriate trainees (male, female, or both); 600 of 1,200 occupations presently taught in vocational-technical schools (which train youth for the most demanding and rewarding blue-collar occupations) are closed to women (Swafford, forthcoming). McAuley (1978) argues that the degree of occupational segregation has actually increased since the thirties. And it is not uncommon to read of "male and female towns"—towns which are plagued by a marked imbalance of the sexes because Soviet planners have built them around a single enormous plant in which (because of occupational segregation) most employees are of one sex.

Demonstrating that occupational segregation exists in the USSR does not, of course, prove that a dual market (in the sense intended by dual-market theorists)

operates there. Nonetheless, it does reveal that the conditions for discrimination are present: the planners who set the wages for all occupations well know which ones are held primarily by women (presumably the second workers in families). Whether they exercise their option in the effort to minimize labor costs is a matter of speculation. But it is clear that future theorizing on sex differences in earnings should take advantage of socialist as well as capitalist experience.

REFERENCES

- Baldwin, Godfrey
1973 *Estimates and Projections of the Population of the USSR, by Age and Sex: 1950 to 2000*. Department of Commerce, Foreign Demographic Analysis Division, Series P-91, No. 23. Washington, D.C.
- Batkaev, P. A.
1972 *Kak organozovana zarabotnaya plata v SSSR (How Wages Are Organized in the USSR)*. Moscow: Znaniye.
- Bibb, Robert and William H. Form
1977 "The effects of industrial, occupational, and sex stratification on wages in blue-collar markets." *Social Forces* 55:974-96.
- Cain, Glen G.
1976 "The challenge of segmented labor market theories to orthodox theory." *Journal of Economic Literature* 14:1215-57.
- Central Statistical Office (Great Britain)
1977 *Annual Abstract of Statistics 1977*. London: pp. 178-9; 261.
- Cerniglia, S. Joseph
1967 *Wages in the USSR, 1950-1966: Construction*. Department of Commerce, Foreign Demographic Analysis Division, Series P-95, No. 63. Washington, D.C.
- Chapman, Janet G.
1977 "Equal pay for equal work?" Pp. 225-42 in Dorothy Atkinson, Alexander Dallin, and Gail Warshofsky Lapidus (eds.), *Women in Russia*. Stanford: Stanford University Press.
- Council of Economic Advisors
1973 "The economic role of women." Pp. 89-112 in *Economic Report of the President*. Washington, D.C.: U.S. Government Printing Office.
- Dodge, Norton T.
1966 *Women in the Soviet Economy*. Baltimore: Johns Hopkins Press.
1977 "Women in the professions." Pp. 205-24 in Dorothy Atkinson, Alexander Dallin, and Gail W. Lapidus (eds.), *Women in Russia*. Stanford: Stanford University Press.
- Featherman, David L. and Robert M. Hauser
1976 "Sexual inequality and socioeconomic achievement in the U.S., 1962-1973." *American Sociological Review* 41:462-83.

- Federal Bureau of Statistics (Switzerland)
1977 *Statistisches Jahrbuch der Schweiz* 1977. Basel. pp. 364-7.
- Feshbach, Murray and Stephen Rapawy
1976 "Soviet population and manpower trends and policies." Pp. 113-54 in *Soviet Economy in a New Perspective*. Washington, D.C.: United States Congress, Joint Economic Committee (October 14).
- Fuchs, Victor R.
1971 "Differences in hourly earnings between men and women." *Monthly Labor Review* 94:9-15.
- Gordon, David M.
1972 *Theories of Poverty and Underemployment*. Lexington: Heath.
- Gordon, Lev and Eduard Klopov
1975 *Man after Work*. Moscow: Progress Publishers.
- Hoffberg, George
1968 *Wages in the USSR, 1950-1966: Health Services*. Department of Commerce, Foreign Demographic Analysis Division, Series P-95, No. 64, Washington, D. C.
1969 *Wages in the USSR, 1950-1967: Education*. Department of Commerce, Foreign Demographic Analysis Division, Series P-95, No. 66. Washington, D.C.
- Institut National de la Statistique et des Etudes Economiques
1977 *Annuaire Statistique de la France* 1977. Paris. p. 573.
- Karapetyan, S. A.
1968a *Modelirovaniye semeynykh dokhodov po materialam perepisi naseleniya* (Modeling Family Incomes with Census Data). An unpublished dissertation for the candidates degree, Rostovskiy-na-donu Institut Narodnogo Khozyaystva.
1968b *Modelirovaniye semeynykh dokhodov po materialam perepisi naseleniya* (a 23-page published summary of the above dissertation). Rostovskiy-na-donu.
- Katz, Zev, Rosemarie Rogers, and Frederic Harned (eds.)
1975 *Handbook of Major Soviet Nationalities*. New York: Free Press.
- Kerlinger, Fred N. and Elazar Pedhazur
1973 *Multiple Regression in Behavioral Research*. New York: Holt, Rinehart, and Winston.
- Kharchev, A. G. and S. I. Golod
1971 *Professional'naya rabota zhenshchin i sem'ya* (The Occupations of Women and the Family). Leningrad: Nauka.
- Korzhenevskiy, I. I.
1965 *Osnovnyye zakonomernosti razvitiya sprosya v SSSR* (Fundamental Regularities in the Growth of Demand in the USSR). Moscow: Ekonomika.
- Kotlyar, A. E. and S. Ya. Turchaninova
1975 *Zanyatost' zhenshchin v proizvodstve* (The Employment of Women in Production). Moscow: Statistika.
- Kunel'skiy, L. E.
1972 *Sotsial'no-ekonomicheskiye problemy zarabotnoy platy* (Socioeconomic Problems of Wages). Moscow: Ekonomika.
- McAuley, Alastair
1977 "The distribution of earnings and incomes in the Soviet Union." *Soviet Studies* 29: 214-37.
1978 "Women's work and wages in the USSR." Paper given at the annual conference of the National Association for Soviet and East European Studies, Cambridge, England.
- Mikhailyuk, V. B.
1970 *Ispol'zovaniye zhenskogo truda v narodnom khozyaystve* (The Use of Female Labor in the National Economy). Moscow.
- Mincer, Jacob and Solomon Polachek
1974 "Family investments in human capital: earnings of women." *Journal of Political Economy* 82:S76-S108.
- Ministère des Affaires Economiques
1976 *Annuaire Statistique de la Belgique*. Brussels. p. 573.
- National Central Bureau of Statistics
1977 *Statistical Abstract of Sweden* 1977. Stockholm. pp. 258-61.
- Netherlands Central Bureau of Statistics
1977 *Statistical Yearbook of the Netherlands* 1976. The Hague. pp. 319-322.
- Oganyan, G. A.
1974 *Raspredelitel'nyye otnosheniya v neproizvodstvenno sfere* (Distributive Relationships in the Nonproductive Sphere). Kiev: Naukova Dumka.
- Rabkina, N. E. and N. M. Rimashevskaya
1972 *Osnovy differentsi yatsii zarabotnoi platy i dokhodov naseleniya* (Principles of the Differentiation of Wages and Income of the Population). Moscow: Ekonomika.
- Rapawy, Stephen
1970 *Wages in the USSR, 1950-1968: Trade*. Department of Commerce, Foreign Demographic Analysis Division, Series P-95, No. 67. Washington, D.C.
- Rimashevskaya, N. M.
1965 *Ekonomicheskii analiz dokhodov rabochikh i sluzhashchikh* (Economic Analysis of the Incomes of Workers and Employees). Moscow: Ekonomika.
- Rzhanitsyna, L.
1971 "Obespecheniye vysokogo urovnya zhizni vsem trudyashchimsya" (Ensuring a high standard of living to all working people). *Sotsialisticheskii trud* 8:11-8.
- Sacks, Michael
1976 *Women's Work in Soviet Russia*. New York: Praeger.
- Sakharova, N. A.
1973 *Optimal'nye vozmozhnosti ispol'zovaniya zhenskogo truda v sfere obshchestvennogo proizvodstva* (The Optimal Use of Women's Labor in the Sphere of Social Production). Kiev: Vishcha Shkola.
- Sarkisyan, G. S. and N. Kuznetsova
1968 *Potrebnosti i dokhod sem'yi* (Family Needs and Incomes). Moscow: Ekonomika.
- Sawhill, Isabel V.
1973 "The economics of discrimination against women: some new findings." *Journal of Human Resources* 8:383-96.

- Schroeder, Gertrude E.
 1972 "An appraisal of Soviet wage and income statistics." Pp. 287-314 in Vladimir G. Treml and John P. Hardt (eds.), *Soviet Economic Statistics*. Durham: Duke University Press.
- Shumilova, Ye. D.
 1964 "Ratsionalizatsiya truda v bytu kak faktor rosta svobodnogo vremeni" (Rationalizing domestic labor as a factor in the increase of free time). *Voprosy filosofii* 12:123-6.
- Statistisches Bundesamt
 1977 *Statistisches Jahrbuch 1977 für die Bundesrepublik Deutschland*. Wiesbaden. pp. 438 and 444.
- Swafford, Michael
 Forth- "The socialization and training of Soviet coming industrial workers." In Arcadius Kahan and Blair Ruble (eds.), *Industrial Labor in the USSR*. New York: Pergamon.
- Szalai, Alexander (ed.)
 1972 *The Use of Time*. Paris: Mouton.
- TsSU (Tsentral'noye Statisticheskoye Upravleniye [Central Statistical Administration])
 1968 *Trud v SSSR (Labor in the USSR)*. Moscow: Statistika.
 1970 *Trud v Armyanskoy SSR (Labor in the Armenian SSR)*. Yerevan: TsSU pri Soviete Ministrov Armyanskoy SSR.
- 1972a *Pol, vozrast, i sostoyaniye v brake naseleniya SSR (Sex, Age and Marital Status of the Population of the USSR)*. Vol. 2 of the 1970 census series. Moscow: Statistika.
- 1972b *Uroven' obrazovaniya naseleniya SSSR (The Educational Level of the Population of the USSR)*. Vol. 3 of the 1970 census series. Moscow: Statistika.
- 1973a *Narodnoye khozyaystvo v 1972 godu (The National Economy in 1972)*. Moscow: Statistika.
- 1973b *Raspredeleniye naseleniya SSSR po zanyatiyam (Distribution of the Population of the USSR by Occupation)*. Vol. 6 of the 1970 census series. Moscow: Statistika.
- 1974 *Migratsiya naseleniya, chislo i sostav semey v SSSR (Migration of the Population; Number and Composition of Families in the USSR)*. Vol. 7 of the 1970 census series. Moscow: Statistika.
- U. S. Department of Commerce
 1977 *Statistical Abstract of the United States: 412*. Washington, D.C.: U.S. Government Printing Office.
- Wiles, Peter
 1974 *Distribution of Income: East and West*. New York: Elsevier.

MANUSCRIPTS FOR THE ASA ROSE SOCIOLOGY SERIES

Manuscripts (100 to 300 typed pages) are solicited for publication in the *ASA Arnold and Caroline Rose Monograph Series*. The Series welcomes a variety of types of sociological work—qualitative or quantitative empirical studies, and theoretical or methodological treatises. An author should submit three copies of a manuscript for consideration to the Series Editor, Professor Robin M. Williams, Jr., Department of Sociology, Cornell University, Ithaca, New York 14853.

THE INFLUENCE OF "FAMILY BACKGROUND" ON INTELLECTUAL ATTAINMENT*

SANDRA SCARR

Yale University

RICHARD A. WEINBERG

University of Minnesota

American Sociological Review 1978, Vol. 43 (October):674-692

"Family background" frequently has been found to have long-term effects on adult intellectual, occupational, and economic outcomes. Since families differ both genetically and environmentally, it has been difficult to interpret family effects in studies of individuals or biological relatives. This study includes samples of adoptive and biologically-related families with children between 16 and 22 years of age. We regressed child IQ on several family demographic variables, on parental IQ, and on natural parent characteristics (for the adopted children) to estimate the degree of genetic bias in the coefficients on measured family background. The results indicate that there is little effect of those family environmental differences studied on IQ differences among the adolescents in the SES range of working to upper middle class. Parent-child and sibling correlations further indicate that genetic differences among families account for the major part of the long-term effects of "family background" on IQ.

Family background has been much discussed and studied recently as a source of inequality among American adults (Behrman et al., 1978; Duncan, 1968; Duncan et al., 1972; Griliches and Mason, 1972; Jencks, 1972; Jencks and Brown, 1978; Sewell and Hauser, 1975; Taubman, 1976; Taubman and Wales, 1972; 1974). That accidents of birth leave us at the mercy of our families' fortunes and that home environments can affect life chances

strike most social scientists as unfair, undemocratic, and even morally wrong. Even more difficult for some to accept is the idea that genetic differences among individuals and families can control some of our differences in adult achievements. The impact of family environmental and genetic differences on intellectual outcome of children is the subject of this study.

It frequently has been reported in recent years that "family background" continues to affect intellectual, educational, occupational, and income differences long after children have grown up and left home. Some vaguely specified characteristics of the offspring are differentially rewarded by employers, and those offspring traits are correlated with parental and home characteristics, even 30 years after the offspring have left home. Although there are substantial differences among studies in the magnitude of the effects they find for family variables (Leibowitz, 1978; Crouse, 1978), there is no sign that the effects diminish with time; in fact, Taubman (1977a) reported stronger effects of "family background" and own IQ as one approaches middle age.

Studies of outcome differences among the offspring of biologically-related families confound four sources of variance: within and between family, envi-

*Address all communications to: Sandra Scarr; Department of Psychology; Yale University; New Haven, CT 06520.

The present study was supported by the William T. Grant Foundation and the National Institute of Child Health and Human Development (HD-08016). The manuscript was prepared while the first author was a Fellow at the Center for Advanced Study in the Behavioral Sciences, with support from the Spencer Foundation and NIMH.

This study was conducted with the full collaboration of the Minnesota State Department of Public Welfare, Adoption Unit, directed by Ruth Weidell and assisted by Marjorie Flowers. Their help was invaluable. The additional support of the Open Door Society, Lutheran Social Service, and The Children's Home Society, all of Minnesota, facilitated the study.

We are very grateful for the assistance of Louise Carter-Saltzman, Harold Grotevant, Margaret Getman, Marsha Sargrad, Patricia Webber, Joanne Bergman, William Thompson, and Carol Peterman in the collection of the data and the preparation of this report. Special thanks go to Arthur R. Goldberger for his extensive help with the data analyses.

ronmental and genetic differences. Regressions of individual outcomes on differences in family background are not illuminating as to the genetic or environmental sources of outcome differences, because parents transmit both genes and family environments, that are likely to be correlated with each other and with genetic differences between families. In other words, the genetic variance in the predictors is likely to be correlated with the genetic variance in the outcomes. As unreconstructed liberals, we get upset about the long-term environmental effects of families on their offsprings' life chances. When individual outcomes are shown to be affected by "family background," we don't know how upset to be.

Behavior Genetic Methods

As Taubman and his collaborators (Behrman et al., 1978) have shown, twin study methods can help to define what is subsumed by the term "family background." Behavior genetic methods have long included the study of genetically and environmentally related and unrelated people (and mice, dogs, etc.). The contrast of effects from similar and different treatments on similar and different genotypes has been a continuing fascination for the field. Families are the usual source of human beings aggregated in related groups. Fortunately, for behavior genetic studies, there are also families who are genetically unrelated but aggregated through adoption. Also beneficial to the field has been the tendency of human populations to produce occasional litters of offspring, some of them genetically identical and others no more alike than sibs. Adoptive families and twins offer unique but different opportunities to study the effects of genes and environments on the outcomes of offspring. The confluence of behavior genetic and social science methods to study the effects of "family background" will provide new insights into true environmental effects.

As Jencks and Brown (1978) have indicated, there are two basic approaches to estimating the importance of environmental differences in determining differences in outcomes. First, they say, one

should begin by offering some meaningful definition of what one means by environment. One strategy is to specify what one means by *measured* environment and study the effects of differences in home background on unrelated children, adopted into the homes. A second strategy is to look at only those environmental influences shared by children reared together. One can estimate the contribution of such influences to phenotypic inequality by calculating the correlation between the phenotypes of genetically-unrelated parents and children and unrelated children reared together.

A third way to obtain an estimate of the true environmental effects of family background would be with identical twins reared apart in uncorrelated environments. Genetic differences would be controlled, while both within- and between-family environmental effects would be free to vary. Unfortunately, child development experts repeatedly have warned about the psychological hazards of giving away one of a pair of twins, and there are simply too few cases, too peculiarly sampled, to make these subjects useful to social science.

Adoptive Families

Adopted children, on the other hand, provide almost as useful data as the rare identical twins reared apart, and they are far more available. Adopted children are not genetically descended from the family of rearing, so that environmental differences between families are not confounded with genetic differences in the children, if the adopted children are randomly placed by adoption agencies. Theoretically, regressions of adopted child outcomes on adoptive family characteristics will provide genetically unbiased estimates of true environmental effects in the population. Unfortunately, adoptive families are selected by agencies for being above average in many virtues including socioeconomic status. Thus, they are always an unrepresentative sample of the population to which one would like to generalize. Although it is possible that the adoptive family coefficients on background are good estimates of the

population values, it is difficult to know without modeling the way in which the families were selected. An easier corrective for the possible bias of selected adoptive families is to have a comparison sample of biologically-related families that are similarly selected.

The study to be reported in this paper includes both adoptive and biologically-related families. The comparison of regression coefficients on measured family background for adoptive families with those of biological families is an estimate of the extent of the genetic bias in studies of family background effects in the usual sociological and psychological studies of families. An additional focus of the paper is on family correlations for IQ. The comparison of similarities among related and unrelated children, and the comparison of related and unrelated parent-child correlations, is the best estimate of the true environmental effects of total, shared family background. The children to be reported here are the oldest adoptive sample ever studied. The study was designed to assess the cumulative impact of family environments at the end of the child-rearing period. If differences in family environments have lasting impact on individual differences in intellectual functioning, the study of adolescents, adopted in the first few months of life, should reveal those differences.

METHOD

Subjects

The 845 subjects in this report are members of 120 biological and 104 adoptive, white families in Minnesota. The adoptive families included 194 adopted and 15 biological children between the ages of 16 and 22. In the first section of this report, only the 150 adopted children whose natural mothers' educational levels were known are included. In the second section, all adoptees are included. The biological families include 237 children with complete data and 268 with IQ data. Adoptive families were recruited through the Department of Public Welfare (DPW), whose director sent letters on behalf of the study to 1,620 families who had adopted children between 1953 and 1959. We were

interested particularly in families who had adopted at least two children, so that our recruitment concentrated on those volunteers with two available children between the ages of 16 and 21 at the time of testing. Table 1 gives the details of adoptive family recruitment.

Of the 1,620 letters sent by DPW, 477 were returned to us without forwarding addresses, which was hardly surprising since the addresses were 15 to 20 years old. Another 345 letters received no response, which may mean that they were not received or that the family chose not to acknowledge our attempted contact, even though follow-up letters also were sent. Of the 798 families known to be eligible to participate, 471 agreed to come to the university for a half-day testing session. Many of those who refused lived far across the state and were unable to join the study. Others did not choose to subject themselves to such extensive scrutiny. The final interview sample who came to the university consisted of 115 families; nearly all of these families have two children in the designated age range and were conveniently accessible to the university. An additional 164 families, most of whom had only one child in the prescribed age range, participated in the mail sample that will not be discussed in this paper. Other willing families were not recruited because of funding and time limitations.

To check on the representativeness of the sample recruited for the study, we

Table 1. Recruitment of Adoptive and Biological Families

Adoptive Families	
Letters sent by DPW	1,620
Letters returned undelivered	477
No Response	345
Eligible to participate	798
Said No	327
Said Yes	471
Participate	
By mail	164
By interview	110*
Biological Families	
Eligible to participate	?
Recruited by adoptive families	41
Recruited by media	153
Participated	
By interview	122*

* The samples reported in this paper.

compared the socioeconomic characteristics of participants and nonparticipants at the time of adoption. Since we had no data on the nonparticipants in later years, this was the best comparison we could manage. There were no age, income, educational, occupational differences between participants and nonparticipants (refusals or nonrespondents) at the time of adoption, but, of course, there may be some current differences in the outcomes of their adoptions or family life histories that we are unable to detect by this method.

The biological families were recruited through newspaper articles and advertisements, word of mouth, and the adoptive families. Approximately 153 biological families came from public media contact and about 41 from recommendations of the adoptive families. Of these, 122 were randomly chosen to come to the university for the full evaluation.

All families who participated in the interview procedure received small payments for their time and transportation and bonuses for recruiting other families. The data were collected from July 1974 to June 1976.

A crucial methodological consideration for any adoption study is the age at which the children are placed with their adopting families. Only early placements can guarantee that potentially confounding, early environmental experiences are minimized. All of the children in this study were in their homes before 12 months of age. Exact age of placement was available for 171 of the 194 adopted children. The mean age of placement into the adopted children's present homes was 2.6 months. Of these 171 children, 109 were placed before two months of age, 158 were placed at or before six months. All but six of the 171 were placed by age nine months. Of all the children for whom placement data were available, there is only one case in which the natural mother may have had social contact during the first 68 days. In all other cases the child left the maternity hospital for the adoptive home or a foster placement. All adopted children were genetically unrelated to their adoptive parents and to each other. The biological children were all full siblings and claimed

to be the biological offspring of both parents tested.

Procedure

Subjects in the sample were administered a three-hour battery of tests and interviews at the University of Minnesota as part of a behavior genetic study of intellectual, personality, and attitudinal similarities within families. The data to be reported here are from the *Wechsler Adult Intelligence Scale* (WAIS; Wechsler, 1955), an individually-administered IQ test. Four subtests of the WAIS were administered: vocabulary, arithmetic, block design, and picture arrangement. The combination of these four subtests has been shown to correlate above .90 with the full scale test score and is generally accepted as a shortened version of the adult test (Doppelt, 1956). The test protocols were scored by an experienced psychometrician who was unaware of the respondents' adoptive status.

After scoring all of the tests, we became aware of a substantial sex difference on three of the four subtests, a fact seldom reported in the literature, but of which the Psychological Corporation seems to have been aware for some time (Herman, 1977). From the point of view of regression analysis, these mean sex differences are not critical, because there are about the same proportion of male and female children in the adoptive and biological samples (47 and 45% male, respectively).

RESULTS

Socioeconomic Variables

The socioeconomic characteristics of the biologically-related and adoptive families are shown in Table 2. Parental educational levels in both kinds of families are .75 to 1 standard deviation above the averages of their cohorts in the population. The occupational prestige of the fathers, rated on the expanded NORC scale (Reiss, 1961) is about 60 in both types of families. Since less than half of the mothers were employed, their occupational ratings were not used in the analyses. Family income averages \$25,000

to \$26,000 in both types of families.¹ The variance of the educational, occupational, and income measures is not as restricted as the high means might imply. In fact, the standard deviations are roughly comparable to the population figures (Taubman's veteran twin sample; Taubman, 1977b). Two points should be made, therefore, about the socioeconomic characteristics of these families: first, the adoptive and biological families are fairly comparable, and second, they both represent selected portions of the SES range in the U.S., both regionally and within the region from which they are drawn. It is well-known that volunteers in social science research are self-selected for better-than-average characteristics of all kinds, and the sample of biological families is at least as biased in SES characteristics as the adoptive one. This is what we hoped would happen, without the statistically hazardous procedure of matching individual families.

The adoptive and biological parents also are comparable in mean IQ scores and in the variance of their scores. Compared with the standardization sample for the WAIS, the fathers are more than a standard deviation above the mean and the mothers about ¾ of an S.D. above. It is not accidental, of course, that samples with above-average income, education, and occupational status also score above the average on a standard IQ test. The standard deviation of the parental IQ scores is only ¾ of that of the population, a significant restriction. Their scores are significantly restricted in range, with the lowest scores in the midnineties.

¹ Occupations of the fathers in the two samples varied from janitor, auto mechanic, small farmer (income < \$10,000), telephone installer, and sheet metal worker at the low end to physician, engineer, college professor, and radio station owner at the high end of the scale. Most occupations were in the middle range of carpenter and printer to insurance agent and building contractor.

The income levels of the families may appear to be higher than they are, unless parental age is taken into account. In 1974, the median family income in the North Central region was \$14,017, but the median family income for families headed by workers aged 45-55 was approximately \$18,000. The families in this sample are less than one standard deviation above that value.

Table 2. Means, Standard Deviations, and Correlations of Adoptive and Biological Family Characteristics

	Biological Children (N = 237)												Mean	S.D.		
	1	2	3	4	5	6	7	8	9							
1 Child's IQ																
2 Father's Education	.10	.26	.24	.10	.22	-.19	-.21	.39	.39				112.82	10.36		
3 Mother's Education	.10	.51	.51	.61	.44	.01	-.36	.56	.24				15.63	2.83		
4 Father's Occupation	.12	.57	.25	.36	.39	.02	-.36	.43	.46				14.68	2.24		
5 Family Income	.06	.50	.40	.46	.47	.01	-.30	.37	.13				62.47	24.73		
6 Birth Rank	-.19	.05	.03	.06	.15	.00	-.25	.38	.19				24,987.34	8,770.43		
7 Family Size	-.05	.04	.11	-.00	.21	.10	.08	-.00	.03				1.62	0.63		
8 Father's IQ	.15	.53	.30	.40	.45	.08	.14	-.30	-.10				3.85	1.48		
9 Mother's IQ	.04	.29	.44	.19	.21	.07	.12	.30	.20				118.02	11.66		
10 Natural Mother's Age	-.10	.04	.03	.12	-.02	-.11	-.04	-.10	.03				113.41	10.46		
11 Natural Mother's Education	.21	.33	.24	.29	.43	.09	.14	.20	.10							
12 Natural Mother's Occupation	.12	-.00	.13	.11	.06	-.06	.11	.11	.15				.33			
Mean	106.19	14.90	13.95	60.30	25,935.00	1.43	2.87	116.53	112.43	9	10	11	11.97	30.44		
S.D.	8.95	3.03	2.06	24.14	10,196.78	0.57	1.20	11.36	10.18	9	10	11	1.66	23.24		
				Adopted Children (N = 150)												

r ≅ .16, p < .05.

The children of the two types of families are quite comparable in age, the mean being about 18½ in both groups. The range of ages is 16 to 22 in both groups (with a few older or younger exceptions). There was no correlation between age and IQ. The IQ scores of the adopted children are about 6½ points lower than those of the biological children, however. These results also are shown in Table 2. If IQ is heritable to any extent, one should expect the biological offspring of bright parents to have higher IQ scores than unselected people. The adopted children are not a genetically selected group. Their natural mothers averaged 12 years of education at a present average age of 41. The median educational level for women, aged 25-44 in the Minnesota area, is 12.5 years of education. Education is an indirect measure of intellectual ability, but as we have shown in another study, there is good reason to expect that intellectual level of the natural mothers is reasonably well indexed by their educational levels (Scarr and Weinberg, 1976;1977a;1977b). Furthermore, there was a large study of unmarried mothers in the state of Minnesota during the years 1948-52, when IQ tests were mandated for all women giving up children for adoption. The average IQ score of 3,600 women was 100.00 with a standard deviation of 15.4 (Pearson and Amacher, 1956). Since our mothers were sampled from 1953-59, there is no reason to expect them to differ significantly from the normal population. Fathers, of course, should not be expected to deviate from the average of the population any more than mothers. Thus, the adopted children are genetically a sample of an intellectually average population, while the biological children are more selected.

Correlations among Parental Characteristics

The parental educational levels, family income, and father's occupation are similarly correlated in the biological and adoptive families. Despite the above-average means on all of these variables, the correlations are either greater or of the same magnitude as those reported from more representative samples by Sewell and

Hauser (1975), Jencks (1972), and others. These two facts—the comparability of correlations in the two samples and their comparability with more representative samples—encouraged us to proceed with the regression analyses.

As Table 2 shows, mothers and fathers in the adoptive and biologically-related families are assortatively mated for educational level with a correlation of about .50. Sewell and Hauser (1975:72) reported .52. Father's education correlated with his own occupational status (NORC scale) about .59. Sewell and Hauser reported .43 (Duncan SEI). Father's occupational status correlated with family income about .46, the same figure obtained by Sewell and Hauser. Mother's education is somewhat more correlated with father's occupational prestige in biological than adoptive families (.36 vs. .25), and Sewell and Hauser reported .29. In these samples, mother's education correlated more highly with family income (.40) than in Sewell and Hauser's study (.24), perhaps because our mostly urban mothers may be more likely to be contributing to that income.

From an examination of the means, variances, and correlations of family demographic characteristics, we concluded that there were no important differences between the adoptive and biological families in the study. The correlational patterns were sufficiently similar to those for more representative samples that the regression analyses are probably more directly generalizable to the general population than we had feared from the selected characteristics of the families.

Parental IQ Correlations

Fathers' and mothers' IQ scores were moderately correlated with the family demographic characteristics, as might be expected. In both the adoptive and biological families, father's IQ was more highly correlated with his educational attainment than mother's was with hers. We suppose this says something about selection for advanced education for women in the cohort that is now 45-55 years of age. Adoptive fathers' correlation of IQ score with occupational prestige is a bit lower

than the biological fathers' (.39 vs. .51). Adoptive parents' IQ scores correlated .31, and biological parents', .24, a moderate difference in assortative mating for IQ. There are no other striking differences in the correlations by family type.

Family Size and Birth Rank

The adoptive families have on the average fewer children than the biological families (2.9 vs. 3.9). The average birth rank of those children who were of appropriate age to participate in the study, however, did not differ much in the two types of families. In both cases, the participants were between first- and second-borns, on the average (1.4 and 1.6 in the adoptive and biological families, respectively). This means that the participants from the biological families have a larger number of younger siblings than the adopted children.

Parental characteristics surprisingly are correlated with family size in the biological families. Although it has often been reported in the general population that family size is negatively correlated with parental IQ, occupational status, education, and income, we did not expect to find such relationships in a socioeconomically advantaged sample. Yet, number of children is significantly negatively correlated with all of the family demographic characteristics and with father's IQ in the biological families. As we did expect, adoptive families with more children (the range of family size was from one to six children), were slightly more advantaged than those with fewer children, presumably because adoptive agencies select parents who can afford to rear more children.

Correlations with Children's IQ Scores

It is clear from Table 2 that parental education, family income, family size, and parental IQ tend to be more highly correlated with biological than adopted adolescents' IQ scores. (Father's occupation and birth rank are not.) The greater resemblance between adolescents' IQ scores and their parents' characteristics in biological families presumably results

largely from the genetic resemblance, since both types of families share the home environment (at least, after the first two months of the child's life). The slight correlation between adopted child IQ and family demographic characteristics is confounded by the selective placement of children of better educated (probably brighter) natural mothers into adoptive families with higher levels of parental education, income, and occupational status. Since natural mother's educational level is moderately correlated with the adopted child's IQ, the correlations between adoptive family demographics and child IQ are inflated by the natural mother-child resemblance via selective placement.

Family size is unrelated to child IQ in adoptive families, but negatively correlated in biological families, probably because of the negative correlation between family size and parental characteristics in the biological families. From the adoption data, however, it is clear that family size per se is not a detriment to IQ in the range of adoptive family sizes represented in this study and at the socioeconomic levels of these families. Birth rank, on the other hand, is clearly related to IQ in both the adoptive and biological families. Later-born or adopted children are at a slight disadvantage in IQ.

Selective Placement

Adoption agencies are not blind. They have information about the natural mothers' educational levels, occupational prestige and age, and they use it to match the children of the natural mothers to adoptive families. As shown in Table 2, there are substantial correlations between natural mothers' educational levels and the adoptive families' demographic characteristics, particularly family income and fathers' education. Fortunately for the study, the agencies do not have information on the IQ levels of the adoptive parents or the natural mothers, so that their effective matching for IQ is quite poor. The correlations of adoptive parents' IQ and natural mothers' education are only .20 and .10 for mother and father, respectively. If the correlation between natural

Table 3. Unstandardized Regression Coefficients of Adolescent IQ on Family Demographic Characteristics and Parental IQ in Biologically-Related and Adoptive Families

	Bio.	Adopt.	Bio.	Adopt.	Bio.	Adopt.	Adopt.	Adopt.
Family Characteristics/N	237	150	237	150	237	150	150	132
Father's Education	.855	*	.795	*	.262	-.153	-.248	-.074
Mother's Education	.551	.362	.465	.343	-.525	.378	.336	.282
Father's Occupation	-.065	.040	-.069	.038	-.059	.032	.035	.014
Family Income	.170	-.020	.160	.010	.100	-.020	-.090	-.085
Birth Rank			-2.699	-3.063	-2.724	-3.078	-3.419	-4.077
Number of Children			-.720	-.303	-.751	-.390	-.528	-.605
Father's IQ					.274	.125	.115	.091
Mother's IQ					.357	-.020	*	-.021
Natural Mother's Education							1.325	1.554
Natural Mother's Age							-.226	-.121
Natural Mother's Occupation								.009
R ²	.107	.019	.145	.059	.309	.075	.138	.157

* $F < .01$, variable did not enter the equation.

mothers' educational and IQ levels is .70, as Jencks (1972) believes, then the average of the correlations between natural mother's and adoptive parent IQ levels is only $(.15)(.70) = .105$. Since the agencies have little or no information about the natural fathers, the correlation between the IQs of natural and adoptive parents is undoubtedly lower than .10. This creates a small shared genetic variance in adoptive families, and accounts for less than 1% of the genetic variance in the population, compared with biologically-related families who share half of the genetic variance.

Regression of Adolescent IQ Scores on Family Characteristics

The major concern of this paper is with the predictability of children's intellectual outcomes from their family's demographic and intellectual characteristics. In the first set of equations, shown in Tables 3 and 4, father's education, occupation, mother's education and family income were used to predict the child's IQ. In the biological families parental education and family income are positive coefficients, and father's occupation is negative. This last, seemingly anomalous, result probably reflects the multicollinearity of the family demographic variables, as shown in Table 2. Once all of these intercorrelated variables are in the regression equation, one or more is likely to be pulled in a negative

direction. More attention, therefore, will be given to the R^2 's than to the particular regression coefficients. The total R^2 for the regression of biological children's IQ scores on their families' demographic characteristics in this sample is .107.

The total R^2 for the adopted child regression on the same variables is much lower, only .019.² The positive coefficients on parental education are lower than those in the biological family regression; family income is slightly negative and father's occupation is moderate and positive.

When birth rank and family size are added to the equation, the R^2 's for both the biological and adoptive children increase by about .04 to .145 for the biological offspring and .059 for the adopted children. (The "birth rank" of the adopted children is their social, sibling order in the adoptive family; nearly all adopted chil-

² It was suggested by one reviewer that the differences between the pairs of regression equations be tested by the Chow test. We have resisted calculating yet another statistic because our goal was magnitude estimation, not testing all possible null hypotheses. More importantly, the unequal sample sizes of the adoptive and biological families yield different expected mean squares, and any result would only be approximately correct, with unknown distributions and standard errors. Two leading textbooks on regression analysis (Cohen and Cohen, and Kerlinger and Pedhazur) either do not mention the test or are opposed to its being done, even with equal sample sizes.

Table 4. Standardized Regression Coefficients of Adolescent IQ on Family Demographic Characteristics and Parental IQ in Biologically-Related and Adoptive Families

	Bio.	Adopt.	Bio.	Adopt.	Bio.	Adopt.	Adopt.	Adopt.
Family Character- istics/N children	237	150	237	150	237	150	150	132
N families	120	104	120	104	120	104	104	99
Father's Education	.233 ^a	*	.217 ^a	*	.072	-.052	-.084	-.025
Mother's Education	.119	.083	.101	.079	-.113	.087	.077	.066
Father's Occupation	-.155 ^a	.108	-.166 ^a	.102	-.140	.085	.094	.037
Family Income	.145	-.027	.139	.015	.089	-.019	-.104	-.099
Birth Rank			-.162 ^a	-.195 ^a	-.166 ^a	-.196 ^a	-.218 ^a	-.253 ^a
Number of Children			-.103	-.041	-.107	-.052	-.071	-.079
Father's IQ					.308 ^a	.158	.146	.116
Mother's IQ					.361 ^a	-.023	*	-.024
Natural Mother's Education							.246 ^a	.293 ^a
Natural Mother's Age							-.147	-.074
Natural Mother's Occupation								.023
R ²	.107	.019	.145	.059	.309	.075	.138	.157
F _{R² > 0}								
(d.f. = # families)	3.44 ^a	0.48	3.19 ^a	1.01	6.20 ^b	0.96	1.49	1.68

* $F < .01$, variable did not enter the equation.

^a $p < .05$.

^b $p < .001$.

dren are firstborn of their natural mothers.) Family size is a larger negative coefficient for biological children's IQ scores than for the adopteds', because family size is negatively correlated with demographic characteristics only in the biological families. The coefficients on the demographic characteristics in biological families are reduced slightly when birth rank and family size are added. In the adopted families, the demographic coefficients also are slightly reduced, except family income which is pulled from slightly positive to slightly negative by the addition of birth rank and family size. Birth rank has a higher coefficient for adopted children's IQ than for biological, thereby demonstrating it is entirely a social effect within families.

The addition of parental IQ scores has dramatically different effects on the regression equations in the biological and adoptive families. First, the R^2 for biological children's IQ scores is doubled to .309, whereas the R^2 for adopteds is increased by only .016, to .075. This striking difference in the overall effect of adding parental IQ to the equation must reflect the genetic contribution of biological parental IQ to their offsprings' IQ scores. There are also striking changes in the

coefficients on biological family demographic characteristics once parental IQ has been added. Father's educational coefficient drops to $\frac{1}{3}$ its former value, and mother's education is pulled to a negative coefficient. The coefficients on father's occupation and family income are reduced. Birth rank and family size coefficients remain virtually unchanged, however. The addition of parental IQ to the adopted children's regression changes the demographic coefficients very little, with the exception that the coefficient on father's education is now slightly negative.

The addition of natural mother's education, age and occupation doubled the R^2 for the adopted children, from .075 to .157. The coefficients on adoptive family demographic characteristics are reduced, reflecting a degree of selective placement, with the exception of family income which is more negative than in the equation without natural mother's characteristics. It is natural mother's education that contributes most to the changes in the equation.

Adding information on the natural mother's educational level, occupation, and age increased the R^2 of adopted children by about .09 over the R^2 with just

family demographic, birth order, and family size information. The final R^2 of about .15 is comparable in size to the R^2 of the biological children equation with family demographic, birth order and family size information (.145).

Conclusions from Regression Analyses

Since the social environment is equally well (or poorly) measured for the biological and adopted children, the impact of direct measures of intellectual functioning for the parents is primarily accounting for the genetic contribution of parents to their biological offspring. In this regard, it is noteworthy that the addition of adoptive parental IQ data to the equation for the adopted children has little impact on the adoptive family demographic coefficients, whereas the demographic coefficients for the biological children are greatly changed. Adding parental IQ scores to the equation for the biological children increases R^2 by .16. Presumably having IQ data for the natural parents of the adopted children would cause a similar increment, even though these parents do not rear their children.

From these regression equations it is evident that significant regression coefficients of child IQ on family variables in studies using only biologically-related parents and children are based largely on genetic variance, as indicated by the different R^2 's for the biological and adoptive families.

FAMILY CORRELATIONS

For the second approach to deciphering the meaning of the term *family background*, we used all of the subjects for whom IQ data were available, regardless of what other information might be missing. Thus, the samples of both adoptive and biological family members are considerably larger, ranging from 270 parent-child pairs in biological families to about 180 pairs in adoptive families. Significance levels for the data have been calculated on the numbers of pairs. A more conservative approach would be to use the number of independently sampled families. Which approach is more defensible is not agreed

upon in the literature, and the reader can consult any table of significance levels for correlation coefficients and Fisher's z formula for the calculation of significance levels based on the number of families. Sample sizes for pairs of family members are given in the middle of Table 5, and sample sizes for families at the top.

By calculating the correlations for related and unrelated family members, we hoped to get an estimate of the degree to which similarity in intellectual outcome is conditioned by similarity in the rearing environment. This entails a comparison of biological and adoptive families and a comparison of parent-child with sibling correlations. Parents and children do not share the same rearing environment, whereas siblings do, regardless of their genetic relatedness.

In an earlier study of young adopted and biologically-related children, we found that parent-child correlations were much greater for the biologically-related pairs (yielding heritability estimates in the range of .4 to .7), but the sibling correlations were quite similarly high for both related and unrelated pairs (Scarr and Weinberg, 1977a; 1977b). We speculated that similarities among these young children were greatly influenced by their families' common rearing environments.

In this sample of late adolescents, we were able to check on the degree of family environmental influence at the end of the child-rearing period. The results for the parent-child pairs are quite similar to the earlier study, whereas those for the siblings are very different. The adopted siblings at the average age of 18½ hardly resemble each other at all.

The evidence for genetic effects is striking in all comparisons of correlations among members of the adoptive and biological families. Even though the scores of both biological and adoptive family members have restricted variance, the coefficients for the biological family pairs usually exceed those of the adoptive family members by a statistically significant amount. As Table 5 shows, in total IQ the biological parent-child pairs, the midparent-child and the child-child pairs are significantly more similar than the adoptive family members. Only in vocab-

Table 5. Correlations among Family Members in Adoptive and Biologically-Related Families (Pearson Coefficients on Standardized Scores by Family Member and Family Type) for Intelligence Test Scales

	Reliability	Biological (120 families)				Adoptive (104 families)			
<i>Child Score</i>	(*)	MO	FA	CH	MP	MO	FA	CH	MP
Total WAIS IQ	(.97)	<u>.41</u>	<u>.40</u>	<u>.35</u>	<u>.52</u>	.09	.16	-.03	.14
<i>Subtests</i>									
Arithmetic	(.79)	<u>.24</u>	<u>.30</u>	<u>.24</u>	<u>.36</u>	-.03	.07	-.03	-.01
Vocabulary	(.94)	.33	<u>.39</u>	.22	<u>.43</u>	.23	.24	.11	.26
Block Design	(.86)	<u>.29</u>	<u>.32</u>	<u>.25</u>	<u>.40</u>	.13	.02	.09	.14
Picture Arrangement	(.66)	<u>.19</u>	.06	.16	<u>.11</u>	-.01	-.04	.04	-.03
— = biological > adoptive correlation, $p < .05$.									
Sample Sizes: Pairs of Family Members									
	Biological				Adoptive				
Children	MO 270	FA 270	CH 168	MP 268	MO 184	FA 175	CH 84	MP 168	
Assortive Mating									
	Biological FA-MO				Adoptive FA-MO				
WAIS IQ	.24				.31				
Arithmetic	.19				-.04				
Vocabulary	.32				.42				
Block Design	.19				.15				
Picture Arrangement	.12				.22				
Sample Size	120				103				

MO = mother-child; FA = father-child; CH = child-child; MP = midparent-child.

* Reliability reported in the WAIS manual for late adolescents.

ulary are the adoptive family members similar at a level different from zero. It is no accident that vocabulary differences are most amenable to social environmental influence. Language is the mode of social exchange among human beings, genetically related or not, so that people who live together develop more similar verbal skills than random members of the population. Other skills are not notably similar among people who live together, unless they are genetically related. It also is not surprising that the skill most amenable to mate selection is vocabulary. Evidently, courting couples spend some time talking to each other, but are not as concerned with other intellectual skills!

From these family correlations one can calculate the differences between the adoptive and biological correlations and, depending upon the model, the heritabilities. Genetically-related persons in ordinary families share about half of their genes. Unrelated people share none of their genes, except through the selective placement of adopted children for IQ, of which there is only a slight bias in this study, as explained earlier. Even though

they have always lived together, the correlations of adoptive fathers' and mothers' IQs with adopted children's IQ scores are .15 and .04, respectively, so that there is little evidence for either selective placement or social environmental influence on IQ differences.

Table 6 gives the difference between the IQ correlations of biological and adoptive relatives and the heritabilities, based on a simple-minded model: multiplying the difference between the correlations of biologically-related and unrelated pairs by 1.6, based on biological families sharing half of the total genetic variance plus that portion due to assortative mating ($r = .25$ for parents). A footnote to the Table explains this calculation. This naive model throws the genotype-environment covariance (if any) into the genetic term, because only biologically-related parents transmit both genes and environments to their offspring. The heritability terms calculated here are really additive genetic variance plus GE covariance in the parent-child comparisons and broad heritability (including some dominance) in the sibling comparisons. The inexactitude

Table 6. Differences between the Correlations of Genetically-Related and Unrelated Family Members and "Heritabilities"

	Related-Unrelated				1.6($r_{mp-c} - r_{ad-c}$)*			
	MO	FA	CH	MP	MO	FA	CH	
Child Score								
Total WAIS IQ	.31	.24	.38	.38	.50	.38	.61	
Subtests								
Arithmetic	.27	.23	.27	.37	.43	.37	.43	
Vocabulary	.10	.15	.11	.17	.16	.24	.18	
Block Design	.16	.30	.16	.26	.26	.48	.26	
Picture Arrangement	.20	.10	.12	.14	.32	.16	.19	

MO = mother-child; FA = father-child; CH = child-child; MP = midparent-child.

* The usual calculation for heritability would be to multiply the difference between the biological and adoptive family correlations by two, because the resemblance of bio members depends on sharing half their genes and home environments and that of adoptive members on sharing only the family environment; thus, the difference equals half of the genetic variance in the populations from which the families were sampled. But biological parents and their children (and siblings) are genetically related by half only when parents are mated randomly for the trait being measured. Because parents are not randomly mated for intelligence (the correlation being about .25 in this sample), there is less genetic variability within the biological families, which leads to a higher correlation among the biological family members. To correct for this in the comparison of biological and adoptive family pairs, it is necessary to multiply the difference between the pairs by 1.6 rather than two based on the following formula:

$$r_{bio} - r_{adopt} = \frac{1+m}{2} h^2$$

where m is the phenotypic correlation between the parents.

of the measures, however, makes this distinction academic, in all probability.

The differences between biological and adoptive family correlations in total IQ range from .24 to .38. Multiplying this difference, then, we find that the values for the combination of genetic variance and GE covariance range from .38 to .61. Although this range of heritability values is a far cry from .80, it is substantially different from zero.

In the simplest-minded genetic model that assumes no environmental transmission or genotype-environment covariation, the regression of offspring value on midparent value is an estimate of narrow heritability or the proportion of additive genetic variance in the total variance (Falconer, 1970). The value of the midparent regression coefficient for total IQ is .52, as

shown in Table 5. By a more sensible model for behavioral traits, one that allows for environmental transmission, the regression of adopted offspring on adoptive midparent values is subtracted from the biological midparent-child regression. The resulting value of the midparent heritability estimate for total IQ is .38 in the population from which we sampled.

We have focused on the total IQ score for several good reasons. First, the other tests are parts of this larger whole. Second, the subtests are less reliable than the total score. And, last, the meaning of the whole is greater than the parts taken singly. It is also clear that total IQ has the highest heritability as estimated from the parent-child correlations and from the sibling comparison. These results lead to the same conclusion reached earlier from the regression of child IQ on the family background and parent IQ data; namely, that half or more of the contribution that parents make to differences in their offsprings' intellectual level is genetic.

We have resisted so far, from ignorance and fear of some formidable critics, the temptation to analyze our data in more sophisticated ways. We cannot defend all of the assumptions that must be made to justify elaborate models, and therefore have hesitated to throw ourselves into an inevitable fray.³ Nonetheless, it seems

³ At the time of writing, Morton and Rao (1977), Cavalli-Sforza and Feldman (1973;1977), and the Birmingham group in genetics (Eaves, 1975;1976; Jinks and Fulker, 1970; Martin and Eaves, 1977) each have proposed various models for the transmission of family effects. Goldberger (1975;1978) has questioned the assumptions and specifications of most of them. There is no one set of assumptions or parameters that is satisfactory to convince unbelievers. Therefore, we have presented our data in a form that can be modeled by the various groups, who may then defend their own models.

An analysis of means from this adoption study in relation to biological and cultural transmission of intellectual skills recently has been done by Cavalli-Sforza and Feldman (1977). Using parental education as an environmental index, they obtain an estimate of cultural transmission (n) of intellectual skills in the adoptive families as follows:

We can obtain an estimate of n from the mean IQ of adopted children (which is $6.2/15 = 0.41$ standard deviations above the general mean of the population):
 $n = 0.41/0.74 = 0.55 \pm 0.06.$

evident to us that the study of adoptive and biological families provides extensive support for the idea that half or more of the long-term effects of "family background" on children's intellectual attainments depend upon genetic, not environmental, transmission. Furthermore, in the range of environments sampled in this study, there is little evidence for any measured environmental effects in "family (SES) background." Birth order is the only variable with substantial effects in the adoptive families, and that accounts for about 4% of the IQ variation among the adolescent children.

DISCUSSION

Accidents of birth do leave us at the genetic mercy of our parents, it seems. Different people have different responses to the same environment, and the effects of differences in environments within the range we sampled are very small. The comparison of the coefficients of child IQ on family background would lead one to conclude that in unrelated families the effects of the demographic variables we measured are nearly nil. Even adding a direct measure of social parental IQ does not substantially increase the explained variance for adopted children's IQ differences.

The IQ coefficients for biologically-related children are highly biased in regression equations, because the demographic variables are indirect measures of the parents' abilities, which are transmitted to the offspring genetically as well. Adding demographic information about one of the natural parents of the adopted children doubles the explained variance, even though that parent has never had social contact with the child after the first few days in the hospital nursery. If we had information about the other parent, there is every reason to believe that the R^2 would rise considerably. Thus, the final

equation for the biologically-related children with an R^2 of .31, is about four times as great as that of the adopted children with comparable information about the social class environment alone (even including some selective placement).

It may be thought by some readers that some unmeasured variables that *really* matter in determining children's intellectual development do not vary in these adoptive families, which were selected by the adoption agencies. To argue that the lack of effect of differences among the demographic and intellectual characteristics of the adoptive families is due to this underlying lack of variation, one must simultaneously explain the considerable regression of child IQ on the same family variables, in the same ranges, in the biologically-related families. Presumably, the argument would be that the biological families were not screened by agencies and do vary on those unmeasured family characteristics that *really* matter.

Fortunately, in a younger sample of transracially-adopted children, we have the same data on adoptive families with their own biological children. Table 7 gives these data. For 143 biological offspring of the adoptive parents, the R^2 from the regression of child IQ (at an average age of ten) on family demographic and parental IQ is .301. For the adopted children *in the same families* ($N = 111$, at an average age of seven), the R^2 is .156, or about half of the coefficient for the biologically-related children. This result is

Table 7. Regressions of Child IQ on Family Demographic Characteristics, and Parental IQ in Transracial Adoptive Families with Their Own Children

	Biological Children (143)		Early Adopted Children (111)	
	B	beta	B	beta
Mother's IQ	.474	.32	.141	.13
Father's IQ	.513	.40	-.028	-.02
Father's Education	.682	.14	.389	.09
Mother's Education	-.943	-.15	1.501	.25
Father's Occupation	-.174	-.23	.008	*
Family Income	.445	.06	-.371	-.06
Total R^2	.301		.156	

* $F < .01$, variable did not enter the equation.

The indication from this preliminary analysis is that the results from means make cultural inheritance about as important as biological inheritance. (Cavalli-Sforza and Feldman, 1977:10)

Their analysis is in agreement with a heritability estimate of .5, the value we propose from our family correlations (Table 6).

in accord with Burks's (1928) regression of adopted and biologically-related children's IQ scores at an average of seven years on family background indicators. She found R^2 's of .37 and .18 for the biological and adoptive families, respectively.

We have argued (Scarr, 1977) that the younger adopted children's intellectual skills are more affected by their parents' characteristics and family environments than the adolescents in the present study, who at the average age of 18½ years, have "gone their own ways" in school and community settings and are less subject to the effects of family differences than are younger children. Nonetheless, the selection of adoptive parents by agencies does not decrease the impact of family differences upon their biological children, and differences among the same parents have less impact on their adopted children.

Compared with the regression equations for biologically-related adolescents, the magnitude and signs of the regression coefficients for young biologically-related children are surprisingly similar. The regressions of biologically-related children on measures of their family background are found to be rather stable across samples and greatly inflated by the shared genetic variance in families.

One could argue that the range of environments sampled here is not sufficiently great to bear the weight of any conclusions about the effects of environmental variation in the population. Our counterargument is twofold. First, the comparison with similarly-sampled biological families reduces the force of the argument. Second, the coefficients of the biological families are much like those in other studies with more representative samples.

Even if differences in several demographic measures of family environments do not contribute much to differences in offspring's IQ scores, however, one must not conclude that the levels of environments in general make no difference for the development of intelligence. Obviously, the average performance level of the adopted children depends on the average value of their environments. In this

sample, the average level of the environments is above average, and so is the average IQ level of the unrelated children. Presumably, if they had been reared in below-average homes, their average IQ levels would also be below average.

The average IQ of 106 for the adopted children can be partially explained by selection and partially by SES advantages. First, children who obviously are damaged or genetically defective are less likely to be placed for adoption. If agencies eliminated from the pool of potential adoptees all of the retarded, possibly 3% of the population with a mean IQ of 60, the average IQ score of the adoptable 97% would be 101. Second, if the actual regression of adopted children's IQ scores on family demographic variables is used to predict IQ improvement, an R of .138 yields 2.1 IQ points. Thus, the adoptees would be predicted to have an IQ average of 103.1, not 106.2, given an SES advantage one standard deviation above the population mean. By the same token, the average IQ scores of the biological children would be predicted from SES alone to be 104.9. With the addition of their genetic advantage, the average IQ of biological family adolescents should be 108.0. This is 4.8 points below their obtained average IQ of 112.8. Where do the extra three to five points come from?

One hypothesis is that SES is not a perfect indicator of the child rearing advantages enjoyed by families who volunteer for social science research; they also are above average in their interest in their children. Since we have no reason to believe that working-class families are on the average less interested in their children's welfare than professional families, volunteers would not bias the *slope* of the SES regression but would affect the intercept. Another hypothesis is that the regression of child IQ on family characteristics is not linear over its entire range. In the range we measured, from working to upper middle class, the slope is relatively flat, but it falls off sharply in the lower SES groups. Based on the obvious negative effects of very impoverished environments on children's development we prefer the latter, although our data will not discriminate the two hypotheses.

The Evidence on Individual Differences as Genes and Environments

From our family studies, the evidence of some genetic individual differences in IQ is simply overwhelming. Especially if one considers the past literature, there are literally dozens of studies that support that mild conclusion. When one attempts to get quantitative about proportions of genetic variance in IQ scores, one has to establish a range of probable values rather than any point estimate. There are several reasons for this. First, there may be real developmental differences in the degree to which environmental influences are potent determinants of individual differences. It seems from limited evidence that younger children may resemble their parents more on environmental grounds, because they are more exclusively influenced by their parents before they are launched into the world of schools, social institutions, and many individual choices.

Second, different cognitive skills that are sampled by different measures, such as vocabulary compared with other skills, may be more or less environmentally influenced. Thus, different age groups using different measures may well get somewhat different results. And, third, there are all the measurement and reliability questions that pertain to any study of cognitive abilities.

Going straight to the heart of the matter, we think that most evidence points to a heritability for IQ of about .4 to .7, given that heritability here means the proportion of variance among individuals sampled in twin and family studies, which, as we have repeatedly noted, are not representative of lower SES, neglectful, or abusive environments. If one could include people with really poor environments, the proportion of environmental variance might rise; on the other hand, the genetic variance also might be increased. It is hard to predict whether the proportions of variance would change or not, and in which direction.

It is important to note the lack of systematic, measured, environmental differences among the adolescents. This suggests that within a range of humane environments, from an SES level of work-

ing to upper middle class, there is little evidence for differential environmental effects. The average level of these environments is such that the children perform intellectually somewhat above the population average, even though they have average biological parents. Thus, the environments sampled in family studies are better than average at fostering intellectual development. But why are the relatively poor families rearing adopted children whose IQ scores are nearly as high as those in professional families? It must be that all of these seeming environmental differences that predict so well the outcome differences among biological children are not primarily *environmental* differences, but indices of genetic differences among the parents and their biological offspring. This brings us to social class.

The Evidence on Social Class Differences as Genes and Environment

In 1938 Barbara Burks compared her California adopted and biological children with those studied by Alice Leahy in Minnesota. Grouping the children by the occupational status of their adoptive families, Burks computed the average effects of being born to and reared by, or only reared by, families at different locations in the social structure. As in all adoption studies, the families do not vary over the whole SES range; in fact adoptive samples always omit those lower portions of the income and educational distributions where big negative effects can occur. Nonetheless, it is interesting to examine the overall effects of being reared by a skilled working-class family, or a white-collar family, or a professional family. As we already know, the intellectual levels of parents in those groups differ on the average. What about the children?

For biological children of these occupational classes, the average difference between working-class and professional families was 12 IQ points in Burks's study and 17 IQ points in Leahy's. Children adopted by families of the same occupational classes, however, differed far less—about five IQ points in both studies. Adopted children in professional families

scored below biological offspring; in working-class families, adoptees scored above the natural children; a very predictable genetic outcome. In our Minnesota studies, we found that the natural children of the transracial adoptive families averaged four to six IQ points above their adopted siblings (Scarr and Weinberg, 1976; 1977a). The adolescent adoptees averaged six IQ points below the biological children of comparably advantaged families. As in the other studies, there is a far greater relationship between parental social class and child IQ in the biological than adoptive families.

Since there is always some selective placement of adopted children into families that resemble their biological parents, the actual effect of differences in this middle to high range of social class environments may be less than the five or six IQ points cited. Again, let us emphasize that none of these studies speak to lower-class, deprived, abusive or any other kind of environmental abominations. We are only saying that in that portion of the SES range where so many studies report intellectual differences among children reared in such circumstances, the differences observed among the children may not be primarily of environmental origin at all. From the older studies, Burks (1938) estimated that genetic differences among the occupational classes account for about $\frac{2}{3}$ to $\frac{3}{4}$ of the average IQ differences among the children born into those classes. Our studies support that conclusion.

If this had been a longitudinal study from the first year of the children's lives to the eighteenth, with detailed observations of the children's environments, the regression coefficients of adolescents' IQ scores on a better set of environmental variables may well have been higher. SES variables are far from perfect indices of children's experiences. Presumably, more of the total variance in adolescent IQ would have been accounted for, if better environmental measures had been available. The effect of such a change would be similar in the adoptive and biological families, since the environments of both were equally represented by the SES measures. Thus, the amount of variance explained by *measured* rather than un-

measured environments might be increased in both kinds of families, but the genetic variance estimated would remain the same.

Why Study Genetic Differences in Behavior?

Some readers may conclude that family research supports pessimistic conclusions. What is left to the systematic environment? (Much of the variance is still unexplained, of course!)

We do not see these research outcomes as pessimistic in the slightest. On the contrary, these family studies permit behavioral scientists and social policy makers to sort out important differences in people's environments. There are three major reasons why behavior genetic studies of families are useful.

The first, and weakest one for social policy, is that we need to gain a fuller understanding of the nature of human behavior. The naive environmentalism of the past three decades locked us into assumptions that are simply untenable, useless, and wrongheaded. The average layman had better intuitions about the nature of human differences than many social scientists purported to have. We have the suspicion, however, that most environmentalists privately explained behavioral differences much as the rest of the population does. But why should we continue to be publicly wrong?

The second reason for behavior genetic studies of families is more "relevant," to use a phrase of the sixties. These studies can and do provide diagnostic clues about the nature of some developmental problems. Just as a good family history in medicine and clinical psychology expresses a concern for individual risks, so tracing family patterns of behaviors affords us a look at human behavior in the making, and often a more optimistic prognosis. So, father was a hyperactive boy; today he is a successful business man. So, when mother was a child, she had a difficult time meeting new people; today she is a respected member of community groups. Social scientists can afford to have more respect for the individual patterns of development that make us differ-

ent from one another. Biological diversity is a fact of life; respect for individual differences comes very much from that biological perspective, and is not a trivial victory.

Third, and most important to us, are the implications for intervention programs. In its baldest form, naive environmentalism has led us into an intervention fallacy. By assuming that all of the variance in behavior was environmentally determined, we have blithely promised a world of change that we have not delivered, at great cost to the participants, the public, and ourselves. The fallacy runs like this: if people who do X without our intervention have more desirable outcomes than people who do not do X, then we should persuade, or compel, all people to do X. This is unwise, because some of the reasons for the naturally-occurring differences between those who do and do not do X are not just environmental differences. Many of these seemingly environmental variations are actually genetic differences or gene-environment correlations. People who are different do things differently.

But here is the most costly part of the intervention fallacy: the erroneous belief that small variations in environments within the "humane range" have meaningfully different outcomes for children. If we observe that professional families take their children to the theatre more often than working-class families, or hang mobiles above their cribs more frequently, some social scientists feel justified in recommending to everyone that they take in plays frequently, rather than play baseball in the back yard, or hang mobiles over the crib, rather than carry the baby about wherever they go. Since these are the child-rearing practices of the professional class, whose children excel at IQ tests and in school, all parents are advised to alter their child-rearing practices to follow suit. *It has not been demonstrated that these variations in child rearing are functionally different in their effects on the children*, and we argue that most humane environments are in fact functionally equivalent. Behavior genetic studies of families can spare us all a homogeneity of

environmental practices, imposed by an "omniscient" professional class.

We can do a better job of designing and implementing effective intervention programs, if we know which variations in the environment make a difference and which ones do not. We can shift our resources to the improvement of those circumstances that have clear, environmentally-deleterious effects on people. Many of these we know: we do not have to do research to know that hunger is not good for children, or that child abuse leaves scars. Most of the worst environments are obviously deleterious. But there are many other marginal and less obvious practices and conditions that we can judge only from sophisticated research on the effects of those environments. So, it is important to know what aspects of the environment have consequences for behavioral differences, and which ones are only apparent variations, based on cultural preferences, genetic differences or on gene-environment correlations. People deserve respect for self-expression and their own modes of child rearing, unless there is clear environmental reason to intervene. Behavior genetic methods will help us to gain a far clearer understanding of which environmental variables to worry about.

But, let us recall that *the average level of our environment is the most important determinant of the level of behavioral development*. Therefore, by providing better schooling, nutrition, health care, psychological services and the like, we can raise the average level of the environment and of behavioral development in the whole population. But some of you will argue that there are real dangers for social policy from research on individual and group differences. We see no necessary connections between the scientific results reported here and any social policy. Science is *not* politics, nor are social policies primarily dependent on scientific evidence, however much we might wish sometimes that they were. Policy matters depend mostly on values, and in this society, many groups compete over the translation of their values into policies.

Frankly, we think such pluralism is healthy, because as scientists we have no

special wisdom in policy matters. Our unique gift to the society is the most objective look we can manage at the nature of the human condition. Hopefully, that information will be noticed and used to improve human lives. As citizens, we can try to be heard, so that our work will have the effects we personally value, but in doing so we must be very careful not to throw away our unique contribution—a set of methods and standards of truthfulness that distinguish us from many other groups.

CONCLUSION

The conclusion that we feel is justified by our data is that intellectual differences among children at the end of the child-rearing period have little to do with environmental differences among families that range from solid working class to upper middle class. These results have important implications for sociological and economic studies of the long range effects of family background on adult achievements. The persistent finding that differences in class background bias adult achievements has been interpreted to mean that differences in family environments during the child-rearing period enhance or impede the intellectual, educational, and occupational achievements of the offspring for a lifetime. From our data, it appears to us that these linkages should be reinterpreted to mean that differences in family background that affect IQ are largely the result of genetic differences among parents, which affect their own status attainments and which are passed on genetically to their offspring, whose status attainments are subsequently affected. The implications of these results are that social scientists should be very wary of interpreting the causes and effects of class differences in studies of biological families. We also should be sensitive to the genetic transmission of family characteristics.

REFERENCES

- Behrman, J., P. Taubman, and T. Wales
1978 "Controlling for the effects of genetics and family environment in equations for school-

- ing and labor market success." In P. Taubman (ed.), *Kinometrics: The Determinants of Economic Success Within and Between Families*. New York: North Holland-Elsevier. In press.
- Burks, B. S.
1928 "The relative influence of nature and nurture upon mental development; a comparative study of foster parent-foster child resemblance and true parent-true child resemblance." *Yearbook of the National Society for the Study of Education* 27:219-316.
- 1938 "On the relative contribution of nature and nurture to average group differences in intelligence." *Proceedings of National Academy of Sciences* 24:276-82.
- Cavalli-Sforza, L. L. and M. W. Feldman
1973 "Cultural versus biological inheritance: phenotypic transmission from parents to children (a theory of the effect of parental phenotypes on children's phenotypes)." *American Journal of Human Genetics* 25:618-37.
- 1977 "The evolution of continuous variation III: joint transmission of genotype, phenotype, and environment." Unpublished paper. Departments of Genetics and Biological Sciences, Stanford University.
- Cohen, J. and P. Cohen
1975 *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Hillsdale: Lawrence Erlbaum.
- Crouse, J.
1978 "Effects of academic ability." In C. Jencks (ed.), *Who Gets Ahead?* New York: Basic Books. In press.
- Doppelt, J. E.
1956 "Estimating the full scale score on the Wechsler Adult Intelligence Scale from scores on four subtests." *Journal of Consulting Psychology* 20:63-6.
- Duncan, O. D.
1968 "Ability and achievement." *Eugenics Quarterly* 15:1-11.
- Duncan, O. D., D. L. Featherman, and B. Duncan
1972 *Socioeconomic Background and Achievement*. New York: Seminar Press.
- Eaves, L. J.
1975 "Testing models for variation in intelligence." *Heredity* 34:132-6.
- 1976 "The effects of cultural transmission on continuous variation." *Heredity* 37:51-7.
- Falconer, C. S.
1970 *Introduction to Quantitative Genetics*. New York: Ronald Press.
- Goldberger A.
1975 "Statistical inference in the great IQ debate." Discussion paper 301-75. Institute for Research on Poverty.
- 1978 "Twin methods: a skeptical view." In P. Taubman (ed.), *Kinometrics: The Determinants of Economic Success Within and Between Families*. New York: North Holland-Elsevier. In press.

- Griliches, Z. and W. M. Mason
1972 "Education, income and ability." *Journal of Political Economy* 80:S74-S103.
- Herman, D.
1977 Personal communication, April.
- Jencks, C.
1972 *Inequality: A Reassessment of the Effects of Family and Schooling in America*. New York: Basic Books.
- Jencks, C., and M. Brown
1978 "Genes and social stratification." In P. Taubman (ed.), *Kinometrics: The Determinants of Economic Success Within and Between Families*. New York: North Holland-Elsevier. In press.
- Jinks, J. L. and D. W. Fulker
1970 "A comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behavior." *Psychological Bulletin* 73:311-49.
- Kerlinger, F. N. and E. J. Pedhazur
1973 *Multiple Regression in Behavioral Research*. New York: Holt, Rinehart and Winston.
- Leibowitz, A.
1978 "Family background and economic success." In P. Taubman (ed.), *Kinometrics: The Determinants of Economic Success Within and Between Families*. New York: North Holland-Elsevier. In press.
- Martin, N. G. and L. J. Eaves
1977 "The genetical analysis of covariance structure." *Heredity* 38:79-95.
- Morton, N. E. and D. C. Rao
1977 "Genetic epidemiology of IQ and sociofamilial mental defect." PGL paper. Population Genetics Laboratory, University of Hawaii.
- Pearson, J. S. and P. L. Amacher
1956 "Intelligence test results and observations of personality disorder among 3,594 unwed mothers in Minnesota." *Journal of Clinical Psychology* 12:16-21.
- Reiss, A. J., Jr.
1961 *Occupations and Social Status*. New York: Free Press.
- Scarr, S.
1977 "Genetic effects on human behavior: recent family studies." A Master Lecture, delivered at the annual meeting of the American Psychological Association, San Francisco.
- Scarr, S. and R. A. Weinberg
1976 "IQ test performance of black children adopted by white families." *American Psychologist* 31:726-39.
- 1977a "Intellectual similarities within families of both adopted and biological children." *Intelligence* 1:170-91.
- 1977b "Nature and nurture strike (out) again." Unpublished paper, Department of Psychology, Yale University.
- Sewell, W. H. and R. M. Hauser
1975 *Education, Occupation, and Earnings: Achievement in the Early Career*. New York: Academic Press.
- Taubman, P.
1976 "The determinants of earnings: genetics, family, and other environment; a study of white, male twins." *American Economic Review* 66:858-70.
- 1977a "Ability, IQ, and earnings." Unpublished manuscript, Department of Economics, University of Pennsylvania, Philadelphia.
- 1977b Personal communication, May.
- Taubman, P. and T. Wales
1972 *Mental Ability and Higher Educational Attainment in the Twentieth Century*. Berkeley: NBER-Carnegie.
- 1974 *Higher Education and Earnings*. New York: McGraw-Hill.
- Wechsler, D.
1955 *Wechsler Adult Intelligence Scale*. New York: Psychological Corporation.

CROSS PRESSURES DURING SOCIALIZATION FOR MEDICINE*

ISRAEL ADLER AND JUDITH T. SHUVAL

Hebrew University of Jerusalem

American Sociological Review 1978, Vol. 43 (October):693-704

Processes of socialization with respect to three principal role components of the competent physician are analyzed by means of turnover tables between contiguous periods observed during a longitudinal study. Summary measures concerning the net turnover, transition probabilities and continuous time transition rates permit an analysis of the overall process of change from a point in time prior to the start of formal socialization all the way to the end of year six when formal socialization ends. Findings show that the process of socialization with respect to the People component is characterized by ambivalence which is expressed in almost balanced positive and negative cross pressures concerning the importance of this component in the role of the competent physician. The process of change with respect to the Status component is an absorbing Markov chain: change is in a positive direction and occurs rather quickly so that by the end of year three virtually all of the population has moved to a positive position on this component and practically no change is evident during subsequent years. With respect to the Science component the data show that pressures are consistently and dominantly negative during the entire process with few counter pressures in a positive direction.

Introduction

The context of socialization for medicine is crosscut by a multitude of pressures and messages transmitted from socializers to students as the latter learn the norms and values of their future profession. Socializing agents are not limited to formal instructors but include a wide variety of significant others with whom medical students have contact: paramedicals, technicians, patients, other students, interns, premedical instructors, clinical teachers and others (Shuval, 1975b). These agents are of varying significance at different times and the salience of their communications may be differentially picked up by students. Nevertheless the large number and variety of socializing agents as well as their very different roles, statuses, backgrounds and orientations imply that the socializing context in which

students find themselves is far from homogeneous and is characterized by an extremely broad array of values and norms (Becker et al., 1969; Bloom, 1971; Fredericks and Mundy, 1976; Hughes, 1966; Merton et al., 1957). Professional socialization is a form of adult socialization and is therefore characterized by many of the latter's special qualities (Brim, 1966; Olmstead, 1969).

Certain aspects of this heterogeneous context may be perceived by students as relatively structured, depending on the substantive area considered. With respect to the overall authority structure, for example, there may be considerable consensus among socializing agents so that a fairly uniform set of messages is communicated and picked up. The same may be true with respect to the hospitalized patient's role. However observation of the subculture of medical education suggests that many of the central norms and values of professional practice are transmitted in a rather ambiguous style which is characterized by conflicting messages, and that there may be a lack of consensus about many of them among different socializers (Shuval, 1975b).

At any period of socialization there exists a multitude of simultaneous micro-processes which impinge on students to change or reinforce existing attitudes.

* Address all communications to: Israel Adler; Medical Sociology Section; Hadassah Medical School; The Hebrew University; Jerusalem, Israel.

The research reported on here was supported by Contract No. 06-704-02 and by Grant No. HS 01857-01 and HS 01857-02 from the National Center for Health Services Research, Health Resources Administration, United States Department of Health, Education and Welfare. It was also supported by a grant from the United States-Israel Binational Science Foundation (BSF), Jerusalem.

These processes are varied in their nature, strength and direction, are often inconsistent and may result in ambiguities for students undergoing socialization. Such counterprocesses sometimes balance each other out in which case no change will occur. In other cases, one set of pressures will dominate such that a net directional change can be observed but the weaker forces, while present in the system, will not show themselves much in the outcome.

The forces operating within the system can be observed either directly by tapping objective features of the system, or by measuring the attitudinal consequences of the change processes on the population. In the latter case the data observed are the result of interaction between the student population and the system and depend on the tenacity of attitudes among subjects, on selective perception processes, on individual modes of resolving ambiguities and on the development of group norms. Looking at the attitudinal consequences of the change process allows us to make inferences about the nature of the system which impinges on the group.

We assume that the process which students experience is inherently continuous; it can be viewed as composed of a series of oscillations between any two observations which are necessarily discrete. It therefore seems important to avoid the fallacy of assuming that change occurs discretely or that between any two observations a person changes at most once. Given the assumption of continuity, methodological precautions need to be taken to avoid this line of thinking.

We have reported earlier that students undergoing professional socialization in Israel perceive the physician's role in terms of three normative role components which define the central qualities needed for competent practice. These have been termed People, Status, and Science components of the occupational role (Shuval, 1975a). The People component concerns qualities which relate to the physicians' interactions with others on both a colleague and client level. The Status component includes traits which focus on the practitioner's place in the organization of the profession with particular reference to

its hierarchical structure. The Science component relates to knowledge, research, and other cognitive areas.¹ Medical students show a remarkable similarity and consistency over time during socialization in the priority they attribute to the three components: the People component is viewed as most important, Status is viewed as lowest in priority, while Science falls between the two (Shuval, 1975a).

It would seem that socialization with respect to these components would not be similar since attitudes and orientations toward them differ among socializing agents, and curricular inputs focusing on them vary in magnitude and intensity.

There appears to be a good deal of ambiguity among socializers with regard to the People component. On the one hand there is an awareness, frequently based on a nostalgically-perceived past, that a more human, person-oriented touch used to prevail in medicine but has disappeared in the context of contemporary practice which has lost the capacity to be concerned with the whole human being. While this feeling is prevalent among many socializers, especially patients and paramedicals but also physicians, it is countered by the many situational pressures of contemporary practice and the manifest orientations of many staff members who view the patient as a case rather than as a person. Observation of the teaching setting suggests that both of these themes are communicated by

¹ The items on which the People component is based are: the ability to understand others' feelings and problems, desire to help others and alleviate suffering, ability to arouse others' confidence, desire to solve community health problems, ability to work with and get along with others, the importance of a warm personality, tolerance and flexibility in relations with others. The items on which the Status component is based are: the desire to carry responsibility, emphasis on external appearance, administrative ability, the importance of prestige and status. The items on which the Science component is based are: knowledge and skill in the natural sciences, scientific curiosity, research ability, originality. Respondents were requested to indicate the importance on a five-point scale of each of these items for practice of the competent health practitioner in their respective fields. The analysis is based on summary scores for each component derived from its subset of items (Shuval, 1975a; Shuval and Adler, 1977).

socializers often by means of symbolic and indirect mechanisms.

It is our impression that the Science component generally is accepted and respected as a central component of medical practice by virtually all socializers. Here there appears to be little ambiguity although this component tends, under certain conditions, to conflict with the People component. When that occurs, the Science component usually gains the upper hand relative to People. On the whole Science is highly respected, viewed as the essential basis of practice, and scientific research is much valued especially in university-affiliated hospitals (Mumford, 1970).

As awareness grows of the organizational basis for contemporary medical practice, the Status component assumes increased importance for students undergoing socialization. We would expect that in Israel, where future physicians almost all expect to practice in a hospital setting, the importance of this component would be communicated in a fairly unambiguous manner and in a positive light. The tradition of the solo practitioner has practically disappeared in Israel and the organization-man in medicine is viewed as the desired role model.

In the light of these considerations we would expect students' orientations to these components to change during socialization in a complex manner in response to the heterogeneous array of socializing messages to which they are exposed. Such change, if explored in depth, is likely to reveal responses to a variety of cross pressures which differ in their strength and direction at different periods during socialization. It is also likely that students will shift back and forth in their orientations to a given component, sometimes with considerable frequency and in what may appear to be an erratic pattern. Our view is that such groping reflects the immaturity and uncertainty characterizing students during certain stages of professional socialization and represents a response to the inconsistencies and cross pressures of the process. This paper attempts to explore some of the cross pressures and mini-processes of change that accompany the evolution of

orientations to the three basic components of practice during professional socialization.

These issues can be examined most fruitfully by means of longitudinal research which permits detailed exploration of change patterns in the same population over time. Such a design provides the unique possibility of making inferences regarding the nature of the transition processes as well as the relative magnitude of cross pressures in terms of the balance and amount of movement between times.

Empirical Procedure

The population under study included all medical students who began their studies in 1969 and who took part in the Israel Study of Socialization for the Health Professions: 168 students at the Hebrew University-Hadassah Medical School and at the Tel Aviv University Medical School. The longitudinal design of this research dictates data collection as follows: in 1969 (T_0), at the end of year one (T_1), at the end of year three (T_3), at the end of year five (T_5), and at the end of year six in 1975 (T_6). The last time represents the end point of formal socialization but one more year of internship is required before students are licensed to practice (Shuval, 1973; 1975a; 1975b).

At all times responses were obtained from the same population to an identical set of 15 items which referred to traits of the practicing physician. Respondents were requested to indicate the importance on a five-point scale of each of these items for competent practice (see fn. 1 for items). Previous analysis of the matrix of intercorrelations of this set of items showed a consistent pattern of clustering at all times in terms of the three basic components of medical practice referred to above: People, Status, and Science oriented traits. The analysis presented here is based on summary scores derived from each of the three subsets of items (Shuval, 1975a; Shuval and Adler, 1977).

The present analysis was carried out on the above variables dichotomized at their respective medians at T_0 , and this cutting point was maintained for all subsequent

times. This procedure assures unambiguous meaning of a positive or a negative attitude on any component at all times.

Methodological Strategies

The analysis attempts to uncover the nature of the total process through its sub-segments; thus it permits us to see general smoothness or discontinuity in specific periods. With this in mind three strategies will be utilized which are all based on the raw turnover tables summarizing attitudes with respect to the same variables measured at contiguous times, to be referred to as the transition matrix. The basic structure of the table is presented in Figure 1 for the first two times measured.

Net Turnover Rate

Davis (1963) proposes the Net Turnover Rate, defined as $\frac{(n_{+-} - n_{-+})}{N}$, to mea-

sure the net shift in a population distribution between two observations. NTO ranges from -1 to $+1$. A positive NTO indicates an increase in the proportion of $+$ attitudes, meaning a net shift in the population in the $+$ direction, whereas a negative NTO means a decrease in the proportion of $+$ responses, i.e., a net negative shift in the population. It can be shown easily that in the case of dichotomous variables, NTO is equivalent to the amount by which the proportion of people at the $+$ value increases between two observations.²

The NTO, being a single summary measure, estimates the result of the multitude of forces and processes operating in the system. However, it furnishes no tools for assessing either the existence or the relative intensities of the counterprocesses which might be operant in the system. NTO also does not enable the researcher

		T_1		
		$+$	$-$	Total
T_0	$+$	n_{++}	n_{+-}	$n_{+}^{(0)}$
	$-$	n_{-+}	n_{--}	$n_{-}^{(0)}$
Total		$n_{+}^{(1)}$	$n_{-}^{(1)}$	N

Figure 1. Schematic Transition Matrix

to isolate the operation of the system at different times since it depends on the marginal distributions at the times at which observations are made and these are a product of a variety of processes which have taken place in previous periods.

Transition Probabilities

In addition to focusing on summary measures which tap the results of the various transition processes in aggregate form, one can focus on the transition probability matrix. An estimate of the probability matrix is obtained by calculating percentages in the transition matrix in the row direction. In such a manner one obtains estimates of the conditional probabilities of being at a given state at time T_2 given that an individual was at a specified state at the time T_1 . Focusing on the features of this probability matrix follows the tradition of the Markov-chain analysis (cf. Kemeny and Snell, 1960) which assumes that the behavior of the system is completely determined (in a probabilistic sense), if one is given the initial distribution of the population and the transition matrices.

Since entries in this matrix add up to one in the row direction, there are but two independent statistics in such a matrix: p_{+-} (which denotes the conditional probability of attaining a $-$ value at T_2 given that one has attained a $+$ value at T_1); and p_{-+} (which has a parallel meaning, but denotes transition in the other direction).

These conditional probabilities will be used to estimate the perceived pressure exerted by the socializing system on the

² This follows from the basic fact that

$$n_{+}^{(2)} = n_{+}^{(1)} + n_{-+} - n_{+-}, \quad (1)$$

from which it directly follows, by rearranging terms and dividing by N , that

$$\frac{n_{+}^{(2)}}{N} - \frac{n_{+}^{(1)}}{N} = \frac{n_{-+} - n_{+-}}{N} = \text{NTO}. \quad (2)$$

group undergoing socialization to change attitudes in a given direction. Such pressures stem from formal and informal inputs of the medical school, values and norms of the student culture as well as inputs from outside the socializing institution. p_{+-} will be interpreted as the pressure on the positive population to change in the negative direction, while p_{-+} will be taken as the measure of the amount of pressure on the negative population to change in the positive direction. These counterpressures are not necessarily equal in their inten-

sities. The ratio $\frac{p_{+-}}{p_{-+}}$ indicates the relative magnitude of pressures in opposite direction. The closer this ratio is to one, the more equal the pressures; its maximum is infinity.

It is of some importance to note that these cross pressures are estimated on an aggregate basis and do not necessarily reflect the pressures experienced by any individual. Different patterns of individual behavior can result in the same group pattern. Therefore caution should be exercised in interpreting our data on an individual level.

The conditional probability overcomes the two problems noted with regard to NTO. Rather than addressing itself only to the counterpressures, it provides separate estimates of the strength of the subprocesses that take place within the period observed. Furthermore, being margin-free, it does not reflect earlier processes which are a function of the initial population distribution. p_{+-} might be found empirically to be much higher than p_{-+} (which implies larger pressure to move in the negative direction), while the NTO might be positive and even large in absolute value. This is due to the fact that a forceful subprocess might be operating on a small subpopulation, such that even if all of this subpopulation has changed in a given period, this movement might be overridden by a less forceful movement of a much larger population.

Neither the NTO nor the conditional probabilities address themselves to the possibility of multiple attitude changes between times observed (Davis, 1963). When this possibility exists the transition

probabilities underestimate the amount of movement and one has no way of estimating the size of this bias.

Even if a single movement between observations is assumed, the fact that not all observations are equally spaced complicates the comparisons between the conditional probabilities across periods. Both NTO and transition probabilities are functions of the time elapsed between observations, and since this time elapsed is not constant in our study, the comparisons proposed in this section become even more problematic.

Continuous Time Approach

The problems facing the researcher concerned with discrete-time analysis have been dealt with extensively in the statistical literature and were brought to the attention of social scientists by Coleman (1964). The continuous time approach assumes that transitions may take place at any instant and this may happen numerous times between any two observations. The probability of making a transition is not a constant, but is a quantity which depends on the time elapsed between two observations. Thus, one estimates quantities such that the probabilities of making a move in a specified direction between two observations which are spaced by dt units of time is the corresponding q multiplied by dt . These will be denoted by q_{+-} and q_{-+} , in parallel to the respective transition probabilities.

The Q matrix is called by Coleman transition rates and by others the infinitesimal generator of the process (Mayer, 1972). Its most appealing interpretation for our purposes is that, under the assumptions of the continuous time model, the respective q 's estimate the mean number of times which a member of the population who has started at a given value will change his attitude between two observations (Singer and Spilerman, 1976). Since the p 's and q 's are mathematically related, we shall consider an overall \bar{q} (ranging between zero and infinity) which is the average of the two conditional q 's weighted by the size of the respective groups at the beginning of the period. Thus \bar{q} measures the

extent of transitional oscillations between two observations.³

The transition rates can be used to estimate the transition probabilities for any period of time. Thus, if empirical observations are made at intervals of two years, one can estimate the transition probabilities for the intervening one-year periods, assuming that the nature of the process does not change within this period. Thus, the continuous time approach allows one to solve the problem of varying lengths of time elapsing between observations and the incomparability of estimated probabilities previously discussed.

Findings

Table 1 displays the percentages of positive responses on each of the components at all times. At T_0 the distribution of responses on each of the components reflects the attempt to dichotomize each component at the population median. While none of the distributions is divided exactly at the median, none is too skewed; in all cases, deviation from 50% stem from technical reasons and no substantive meaning should be attributed to them.

From T_1 on, we observe consistent trends in each of the components. The People component displays a negative change over time which expresses a gradual decrease in the importance attributed to the component during socialization. A similar consistent trend is found in the Science component whose proportion of positive responses also decreases, while an opposite trend is displayed by the Status component whose rate of positive responses increases through time.

Due to the dichotomization procedures, caution should be exerted in attributing meaning to the size of these shifts. It is worth noting, however, that in the People component one relatively major shift in the distribution occurs between T_1 and T_3 , when the distribution stabilizes. The Status component displays sizable shifts both between T_0 and T_1 and between T_1 and T_3 , at which time the distribution

Table 1. Positive Responses on Three Components of the Role of The Competent Physician at Five Times (Percentages)

Component *	Year of Observation				
	T_0	T_1	T_2	T_3	T_4
People	58	57	43	45	42
Status	53	74	97	98	98
Science	46	29	18	12	11
N	163	165	157	147	153

* See fn. 1 for items on which components are based. Data consist of summary scores derived from each subset of items dichotomized at its T_0 median.

stabilizes. The Science component differs from the other two in that the shift goes on until T_3 at which time the distribution stabilizes.

Table 2 presents net turnover rates (NTO), transition probabilities (p), and transition rates (q) for all one-year transitions. Within each of the sections of the Table, referring to each of the components, rows portray the parameters of transition between two consecutive years. At times when observations were not made, Coleman's continuous time model was used to estimate intermediary times. Starred entries (*) refer to estimates of the above parameters obtained from a two year interval statistic (Coleman, 1964; Mayer, 1972).

Our strategy in looking at these findings is to seek patterns in the data rather than statistically significant differences between various entries. This approach is governed by our concern with systematic variation of numerous measures estimated at several points in time. While it is possible to test for the significance of differences between NTOs and conditional probabilities, these tests are limited to comparisons of pairs of statistics while our interest is in the systematic patterning of a whole array of such statistics.

Change within Components

People. The NTOs show a small negative balance of movement on the People component between all times but the only transition at which this reaches substantial dimension is between T_1 and T_3 , i.e., the

³ \bar{q} is an approximate statistic since it does not take into account the continuous change in the sizes of the groups.

Table 2. Parameters of Transition Processes

Component	Period of Transition	Net Turn-over NTO	Transition Probabilities			Transition Rates		
			p_{+-}	p_{-+}	$\frac{p_{+-}}{p_{-+}}$	q_{+-}	q_{-+}	\bar{q}
People	T_0-T_1	-.0256	.3587	.4531	.7917	.7380	.9323	.8191
	$T_1-T_2^*$	-.1438	.2521	.1274	1.9788	.3183	.1603	.2503
	$T_2-T_3^*$.2521	.1274	1.9788	.3183	.1603	.2503
	$T_3-T_4^*$.2549	.2014	1.2656	.3403	.2689	.2989
	$T_4-T_5^*$	-.0074	.2549	.2014	1.2656	.3403	.2689	.2989
	T_5-T_6	-.0507	.4154	.2740	1.5161	.7045	.4646	.5739
Status	T_0-T_1	.2372	.1125	.6053	.1859	.1983	.5334	.3586
	$T_1-T_2^*$.2260	.0000	.6373	.0000	.0000	1.0141	.2520
	$T_2-T_3^*$.0000	.6373	.0000	.0000	1.0141	.2520
	$T_3-T_4^*$		—	—	—	—	—	—
	$T_4-T_5^*$.0000	—	—	—	—	—	—
	T_5-T_6	.0000	.0147	1.0000	.0147	—	—	—
Science	T_0-T_1	-.1474	.4648	.1176	3.9524	.6969	.1764	.4159
	$T_1-T_2^*$	-.1096	.4037	.0627	6.4386	.5436	.0845	.2236
	$T_2-T_3^*$.4037	.0627	6.4386	.5436	.0845	.2236
	$T_3-T_4^*$.6651	.0720	9.2375	1.2057	.1305	.2827
	$T_4-T_5^*$	-.0815	.6651	.0720	9.2375	1.2057	.1305	.2827
	T_5-T_6	-.0290	.6471	.0579	11.1762	1.1203	.1002	.2251

* For these periods transition probabilities for one-year intervals were interpolated from observed two-year transitions by the continuous time model (Coleman, 1964; Mayer, 1972).

beginning and end of the preclinical period.

The cross pressures as estimated by the transition probabilities and their ratios confirm that the strongest balance of pressure away from an emphasis on the People component occurs between T_1 and T_3 , the preclinical period; this is shown by the ratios of $1.98 \left(\frac{p_{+-}}{p_{-+}} \right)$ which are the highest

seen on this component. Nevertheless the p 's make it clear that it is the ratio of the cross pressures that is high during the preclinical period, but not their absolute size. In other words this period is characterized by lack of symmetry in the balance of pressures concerning the People component; there is considerably more pressure against People than for it. Indeed there is more absolute negative pressure (p_{+-}) away from the People component at all the other periods, but the counter balance of pressures toward the People component (p_{-+}) is also larger at all other periods so that they more closely balance each other out and their ratios are consequently lower.

In fact the transition probabilities indicate that, with the exception of the preclinical period, the process of socialization with respect to People is characterized by fairly balanced simultaneous pressure toward and away from the People component. Although in most periods the negative pressures are slightly greater, the ratio between them is not far from one. Such a pattern suggests ambiguity on the part of students who reflect almost equal conflicting pressures for and against the importance of the People component in the professional role of the physician. This configuration is of some significance in view of the consistent centrality and priority attributed to the People component by students during all stages of socialization (Shuval, 1975a; Shuval and Adler, 1977).⁴

Looking at the transition probabilities vertically we see that absolute pressures

⁴ Between T_0 and T_1 we find that $p_{-+} > p_{+-}$; nevertheless NTO shows a small negative value. This apparent contradiction reflects the initial distribution of the population on the People component and the dependence of NTO on it.

with respect to the People component in both a positive and a negative direction are greatest at the beginning and at the end of the process. This is also seen if we observe the mean number of moves as expressed by \bar{q} which is largest between T_0 and T_1 and between T_5 and T_6 . The ambiguity felt by students with respect to the People component is apparently greatest at the beginning and at the end of formal socialization when considerable oscillation occurs in their attitudes. During the first year of socialization students who are predominantly positive in their orientation to People find themselves in a curricular environment which is almost purely science-oriented. Apparently this is reflected in uncertainty and in groping with respect to the importance they attribute to People during this time, but the balance of movement suggests that there is still strong pressure favoring People at this time. This positive pressure weakens during subsequent socialization when the science ethic is proportionately more prominent, but it becomes stronger again toward the end of formal socialization when students are crystallizing their image of the competent practitioner.

Status. The NTOs on Status are positive and large up to the end of the preclinical period. No movement is evident subsequently because the entire population moved into the positive category with respect to Status as early as T_3 .⁵ This pattern will be recalled from Table 1.

The transition probabilities show an absorbing Markov chain (Kemeny and Snell, 1960) which indicates strong pressures toward an emphasis on Status with virtually no counterpressure in the opposite direction. The ratios $\frac{p_{+-}}{p_{-+}}$ are therefore

close to zero at all times. The only noticeable departure from this pattern occurs during the premedical period when the polarization of pressures is less extreme than in the other periods. During this time p_{+-} attains its highest value on this component while p_{-+} which is already

large, attains its lowest value. This balance is reflected in the ratio which attains its highest value.

Again we see that \bar{q} is largest during the earliest period of socialization, indicating the existence of relatively forceful cross pressures and oscillation of students between a positive and a negative position.⁶

Science. The NTOs, which show the overall result of various conflicting pressures, indicate a consistent negative trend which is fairly strong during the early stages but which weakens over time and virtually disappears at the end of the process. This means that the total trend may be described as a decelerating negative movement which shows students to be shifting away from an emphasis on the importance of Science but at a decreasing rate over time.

The nature of the cross pressures are revealed more sensitively by the transition probabilities which reflect systemic effects and are margin-free. These probabilities indicate a strengthening of the pressure to deemphasize Science (p_{+-}) precisely during the clinical period. This could be a result of a growing awareness of the uncertainty inherent in much of medical practice and a gradual realization among students that the initial importance attributed to the Science component was unrealistically high, since its place in the physician's normative role, while central, is less dominant than originally thought. This is seen by the lower level of transition probabilities up to T_3 as compared to the higher values after T_3 . The process is described more completely as follows: at the time of entry (T_0) and during the preclinical period when the population is largely positive in its orientation to Science, relatively weak pressures are sufficient to move large numbers to a negative position; later during the clinical period, although the pressures become stronger,

⁶ Since the population is totally immobile on Status after T_3 , the parameters of the continuous time model are not estimable after this time. Therefore the one-year transitions $T_3 - T_4$ and $T_4 - T_5$ also were not estimated.

⁵ Dichotomization of the variables results in some artificiality in this conclusion, but observation of the full range of scores shows the pattern described to be essentially correct.

Note: n_{++} is the number of people who were + at both T_0 and T_1 , n_{+-} is the number of people who were + at T_0 and - at T_1 , and so on.

they are operating on the remaining positively oriented subpopulation which is smaller and therefore the result is a smaller turnover.

In observing the simultaneous pressures for and against the importance of Science in the physician's role, we see that the positive forces are at all times weaker than the negative forces. The only period during which there is any noticeable pressure toward Science is during the premedical period (T_0 - T_1). This is the period during which students are studying under the auspices of the faculty of science at a campus which is geographically separate and distant from the medical school. After the move to the medical school campus, pressures toward Science virtually disappear. Indeed \bar{q} indicates a large amount of movement between T_0 and T_1 as opposed to all other periods. This is apparently a result of the fact that this is the only period during socialization when there are any sizeable cross pressures with respect to Science.

The overall balance of these cross pressures is summarized in the $\frac{p_{+-}}{p_{-+}}$ which shows a sharp monotonic increase over time indicating a change in the relative weight of the cross pressures over time. During the premedical period the ratio is roughly four to one in favor of a negative emphasis on Science, but this increases radically to a ratio of over eleven to one at the end of formal socialization.

Patterns of Change across Components

All components show wide variability on \bar{q} . This indicates that the nature of the process varies in different stages of socialization, i.e., is not time-homogeneous.

Comparing the overall amount of movement across components (\bar{q}) shows a consistent pattern in which the premedical period (T_0 - T_1) is characterized by the largest amount of oscillation and groping. In all cases \bar{q} is largest within the respective components during that period. Apparently this pattern reflects the structure, curriculum and ambience of year one of medical school which is different in sev-

eral respects from subsequent periods of socialization.

During this early period the Jerusalem students were located at a separate campus, the buildings of which were spread in different parts of the city at the time of the study. In both medical schools the program of study was under the auspices and authority of the Faculty of Science which determined not only the content of courses but provided teachers, determined examination procedures, and maintained a powerful stance which affected the climate of student culture. Many students stated that they felt isolated and frustrated in their detachment from the medical school to which they had gained admission after severe competition. It also is worth noting that the curriculum during year one is composed almost entirely of hard science subjects: biology, chemistry, mathematics, physics and a small unit of psychology. If we recall that at T_0 students placed strongest emphasis on the importance of the People component of the competent physician, it is likely that such a curricular atmosphere is in some respects incongruent with expectations. Finally it should be noted that entry into the university represents a sharp shift in students' life styles. Many were living away from home for the first time and all were exposed for the first time to campus life. All of the above factors apparently combine to make this year of socialization relatively anomic and to promote ambivalence and groping.

Ambiguity during year one as estimated by \bar{q} is greatest with respect to the people component, least with respect to Status, while Science falls between the two in its level of ambiguity. It is of interest to recall that the same order among components consistently was found with respect to the importance attributed to them in the role of the competent physician (Shuval and Adler, 1977). It may be suggested that ambivalence is greatest precisely with respect to the component which is most emphasized. When students view a component as very important, they are willing to experiment more with different attitudes concerning it and will therefore be more likely to shift around until they crystallize their position.

The ratio $\frac{P_{+-}}{P_{-+}}$, when viewed across the components, reflects the general atmosphere of socialization by providing a picture of the relative balance of cross pressures with respect to the three components. These differ dramatically from each other in the absolute level of cross pressures and in their patterns of change over time.

The positive and negative forces focusing on the People component are almost balanced during the full course of formal socialization. This is seen by the closeness of the ratio to one. Although the negative pressure is generally somewhat stronger, we have suggested that ambivalence and uncertainty are the probable effects of such a balance. Students are apparently picking up a variety of mixed messages with respect to People which may reflect the uncertainty of socializing agents and possibly of broad segments of the profession as to the place of this component in the physician's role. In any case the data indicate that no clear, dominant viewpoint comes across with respect to this component.

A rather different pattern characterizes the Science component where the balance of pressures is constantly and unambiguously against an emphasis on the component. Here the message is expressed dramatically by the growing size of $\frac{(p_{+-})}{(p_{-+})}$.

The negative pressure (p^{+-}) starts at a moderate level during year one, although, as noted, it has a massive effect then. However, it builds up with decreasing opposition as students move through the years and is powerfully focused toward the end on the small subgroup which has not yet shifted to a negative position on this component.

Pressures with respect to Status are strongly positive and their impact is so powerful that by the end of year three virtually the entire population has shifted to a positive position on this component. So strong is this pattern that there is virtually no negative pressure on students with respect to this component.

In sum the atmosphere in the medical schools may be described as ambivalent

with respect to People, increasingly negative to Science and strongly positive concerning Status.

Summary

The process of socialization with respect to the three principal role components of the competent physician has been analyzed by means of turnover tables between contiguous periods observed. Summary measures concerning the net turnover, transition probabilities and continuous time transition rates permit an analysis of the overall process of change in some detail from a point in time prior to the start of formal socialization all the way to the end of year six when formal socialization ends.

This analysis shows that the process of socialization with respect to the People component is characterized by considerable ambivalence which is expressed in almost balanced positive and negative cross pressures concerning the importance of this component in the role of the competent physician. Although the overall balance of pressure during the entire period tends to be somewhat stronger in the negative direction and these negative pressures show their strongest effect during the preclinical period, most of the other periods are characterized by almost balanced pressures with respect to this component. The first and last years of formal socialization are characterized by the greatest amount of ambivalence with opposing balanced pressures expressing themselves most strongly. We have suggested that this empirical picture reflects the uncertainty and ambiguity of socializers and possibly of the profession with respect to this aspect of the physician's role. Students perceive this lack of a clear message throughout socialization but especially at the beginning when they first encounter the subculture of science and at the end when they feel they are on the brink of practice. This polarization of ambiguity—at the beginning and at the end of socialization—highlights the dilemma of socialization with respect to the People component and suggests that despite students' awareness of its centrality in practice, they feel buffeted by socializ-

ers and ambivalent right up to the end of their formal training when messages from socializers do not become any clearer.

We have described the process of change with respect to the Status component as an absorbing Markov chain. Change is in a positive direction and occurs rather quickly. By the end of year three virtually all of the population has moved to a positive position on this component and practically no change is evident during subsequent years. Since there is no formal curricular input that relates directly to this aspect of practice, it would appear that latent messages with respect to Status are picked up by students. These are almost all positive and are expressed in students' feelings that the organizational aspects of practice are important and should be emphasized. Although on an absolute level students rate the importance of the Status component lower than the other two, their attitude changes in a positive direction during the early part of socialization, apparently in response to certain fairly clear-cut messages that are picked up.

A rather different pattern of change is seen with respect to the Science component. Here the pressures during socialization are consistently and dominantly negative during the entire process with few counter pressures in a positive direction. Despite a strongly science oriented curriculum, the pattern of movement suggests a boomerang effect with reference to the formal curriculum, since movement is toward a decreasing emphasis on the importance of the Science component for the competent physician. These pressures are moderate during the early stages but they are sufficient to move large numbers of students away from the Science component; they increase in strength during the clinical period when they cause most of the diehards who still place a positive emphasis on Science to shift their attitude. One possible explanation could be students' growing awareness of the inherent uncertainty of clinical medicine in which science is central and not entirely dependable.

In sum, this analysis leads to the conclusion that socialization with respect to People is ambiguous especially at the be-

ginning and at the end of formal training, is positive with respect to Status during the early stages and does not backtrack from that position, and is consistently negative with respect to the Science component of practice.

Year one, which is taught under the auspices of the Faculty of Science during which the students have virtually no physical or symbolic contact with the medical school, proves to be a period of groping and uncertainty. Cross pressures are greatest and the data indicate more movement back and forth on all three components during this first year than during any subsequent period during formal socialization. We have suggested that the isolation of students from their home faculty as well as the nature of the curriculum have combined to make this period problematic and to highlight the cross pressures characterizing it. These are especially strong with respect to the People component but they appear with reference to the Science and Status components as well. We have noted that the level of oscillation and cross pressures corresponds directly to the absolute emphasis placed by students on the importance of the components: most on People, least on Status, with Science falling between the two.

REFERENCES

- Becker, Howard S., B. Geer, E. C. Hughes, and A. L. Strauss
1969 *Boys in White*. Chicago: University of Chicago Press.
- Bloom, Samuel
1971 "The medical school as a social system." *Milbank Memorial Fund Quarterly* 49:1-191.
- Brim, O. G., Jr.
1966 "Socialization through the life cycle." Pp. 1-49 in O. G. Brim and S. Wheeler (eds.), *Socialization after Childhood*. New York: Wiley.
- Coleman, James S.
1964 *Introduction to Mathematical Sociology*. New York: Free Press.
- Davis, James A.
1963 *Panel Analysis: Techniques and Concepts in the Interpretation of Repeated Measurements*. Mimeo. Chicago: National Opinion Research Center.
- Fredericks, Marcel A. and Paul Mundy
1976 *The Making of a Physician*. Chicago: Loyola University Press.

- Hughes, E. C.
1966 "The making of a physician: general statement of ideas and problems." Pp. 96-9 in W. R. Scott and E. H. Volkart (eds.), *Medical Care*. New York: Wiley.
- Kemeny, John G. and J. L. Snell
1960 *Finite Markov Chains*. Princeton: Van Nostrand.
- Mayer, Thomas F.
1972 "Models of intergenerational mobility." Pp. 308-57 in J. Berger, M. Zelditch, Jr. and B. Anderson (eds.), *Sociological Theories in Progress*, Vol. 2. Boston: Houghton Mifflin.
- Merton, Robert K., G. Reader and P. L. Kendall
1957 *The Student Physician*. Cambridge, Ma.: Harvard University Press.
- Mumford, Emily
1970 *From Students to Physicians*. Cambridge, Ma.: Harvard University Press.
- Olmsted, Ann and M. Paget
1969 "Some theoretical issues in professional socialization." *Journal of Medical Education* 44:663-9.
- Shuval, Judith T.
1973 "Manpower pools for three health professions in Israel." *Social Science and Medicine* 7:893-910.
- 1975a "Socialization of health professionals in Israel: early sources of congruence and differentiation." *Journal of Medical Education* 50:443-57.
- 1975b "From 'boy' to 'colleague': processes of role transformation in professional socialization." *Social Science and Medicine* 9:413-20.
- Shuval, Judith and Israel Adler
1977 "Processes of continuity and change during socialization for medicine in Israel." *Health and Social Behaviour* 18:112-24.
- Singer, Burton and Seymour Spilerman
1976 "Representation of social processes by Markov models." *American Journal of Sociology* 82:1-54.

STRATIFICATION IN A DUAL ECONOMY: A SECTORAL MODEL OF EARNINGS DETERMINATION*

E. M. BECK, PATRICK M. HORAN AND CHARLES M. TOLBERT II

University of Georgia

American Sociological Review 1978, Vol. 43 (October):704-720

We examine the importance of industrial sectors as hypothesized by the dual economy literature on the process of earnings determination. Using NORC survey data, we find substantively and statistically significant differences in labor force composition and economic status between core and periphery industrial sectors. Our application of a covariance regression model to these data demonstrates that the sectoral differentials in earnings cannot be explained away by differences in labor force composition and that there exist significant sectoral variations in the way that worker characteristics are rewarded. The findings are interpreted as evidence of the importance of structural factors for earnings determination. The critical implications of such findings for neoclassical research traditions in social stratification are noted.

MODELS OF STRATIFICATION AND EARNINGS

Much recent American stratification research has been characterized by an individualistic conception of the relationship

between labor force participation and social rewards. This approach rests on the belief that workers are placed within the socioeconomic order through a competitive process in which skills and abilities of differing value and scarcity are carefully identified, evaluated, and matched with societal needs (Pease et al., 1970). In the study of occupational placement, this conception derives from a functional theory of stratification which posits a competitive matching between the functional importance of occupational roles and the skills and training of job seekers (Parsons, 1940; Davis and Moore, 1945). This theory provides the intellectual un-

* Address all communications to: E. M. Beck; Department of Sociology; University of Georgia; Athens, GA 30602.

Partial research support was provided through G.A.E.S. participation in U.S.D.A. Regional Project S-114. We acknowledge the comments on an earlier draft made by M. Betz, S. Bianchi, R. Cohn, W. Falk, J. H. Madans, M. Killian, P. Farnworth, R. Stolzenberg, and W. Villemez. We especially thank the referees for their constructive critiques.

derpinnings for research which represents occupations as prestige levels and which interprets occupational prestige as a measure of individual success or failure in an attainment process.

Much of the recent analysis by economists of earnings determination likewise has been characterized by an individualistic conception derived from neoclassical economic theories.¹ In this, the human capital perspective, the rational worker invests in training which will maximize the economic return (earnings) on investments (Becker, 1964; Mincer, 1974) while free competition among firms for labor skills guarantees a price for various labor types which is equal to the marginal productivity of that labor.²

Others have noted the similarities between the functional theory of occupations and neoclassical economic theories (see e.g., Bibb and Form, 1977; Grandjean, 1975; Horan, 1978). The point of interest here is less the similarities in the theories themselves than similarities in their implications for empirical research. These neoclassical theories of social structure include both individual participants in the socioeconomic attainment process and a structural context consisting of a set of interrelated employers and firms. The structural context, however, is such that when it is assumed to be working according to the theoretical specifications, it need not be included in analyses. Like Adam Smith's "invisible hand" the competitive structure presumed by neoclassical theory guarantees that differential placement in the socioeconomic order is accomplished in a manner such that this placement is a reflection of a worker's basic value to the system. From this perspective, inequality of earnings must be a reflection of the dispersion of individual resources; low prestige and poverty wages must be the result of resource insufficiency; sporadic employment and job instability must be the products of inade-

quate commitment to work or a weak achievement motivation. In short, socioeconomic success or failure is tied directly to the characteristics brought into the marketplace by the individual workers.

Research deriving from neoclassical conceptions is individualistic in the sense that its explanatory variables tend to be restricted to personal attributes such as formal schooling, work experience, or family background characteristics. Recent attempts to broaden the scope of models of earnings determination (Stolzenberg, 1975; Althausen and Kalleberg, 1977; Osterman, 1975; Kalachek and Raines, 1976; Hudis and Kalleberg, 1977; Bibb and Form, 1977; Spilerman, 1977; Hodson, 1977) have done so by adding to these individual characteristics some supraindividual variables relating to the structure of the earnings determination process—a trend which we will continue in the analysis below.

Although working primarily within the neoclassical framework, Stolzenberg (1975) suggests that labor market competition is not completely uninhibited but is segmented according to occupation. He argues that because of the occupation-specific nature of worker "investments" in preparation for specialized jobs, such training produces "... a strain toward segmentation of labor market competition along occupational lines" (Stolzenberg, 1975:647). In addition to occupational wage effects, Stolzenberg's analysis of census data for employed males also reports substantial industry effects on wages which are interpreted in terms of interindustry variation in the supply and demand for the factors of production.

Osterman (1975) also has investigated the economic implications of segmentation along occupational lines. Using data from the 1967 Survey of Economic Opportunity, he finds that human capital variables tend to explain variations in earnings among jobs in the "primary segment," but not among jobs in the "secondary segment." Analyzing data from the National Longitudinal Surveys, Kalachek and Raines (1976) find that while human capital variables play a major role in the determination of earnings, occupational and industrial factors have an important effect

¹ Much of the sociological research on earnings determination also has been individualistic in orientation and execution. The recent efforts of Hauser and Featherman (1977), Sewell and Hauser (1975), and Duncan et al. (1972) are in this tradition.

² See Thurow (1975:211–30) for a detailed discussion of marginal productivity theories.

on the rate of return on the "human capital stock."

Bibb and Form (1977) juxtapose the human capital model of earnings determination with a "structural" model which includes information on industrial and occupational factors as well as characteristics of the labor force. Analyzing data on the earnings of full-time blue-collar workers, they use regression procedures to compare the explanatory power of human capital variables with structural variables pertaining to industrial and occupational characteristics. They report that their final model accounts for nearly three times as much variance in blue-collar earnings as their human capital model alone.

Stolzenberg (1975), Bibb and Form (1977), and Kalachek and Raines (1976) provide evidence on the importance of organizational characteristics for individual earnings determination, and Bibb and Form raise some important questions about the empirical adequacy of individualistic models. However, the need remains to forge a systematic link between structural theories of earnings determination and the empirical analysis of earnings data. In particular, the nature of the relationship between individual factors and structural factors warrants further consideration. Neoclassical approaches to income determination presume that worker characteristics are paramount in this process and that the income returns to individual characteristics remain fixed across structural settings. While others have addressed the issue of the dominance of individual characteristics, the present analysis will focus on the issue of fixed returns—the assumption that economic returns to worker characteristics are uniform.

SECTORAL MODELS OF EARNINGS

With this in mind, we turn to models of sectoral economic differentiation derived from theories of economic dualism. Such theories encompass a diverse set of literature, diverse both in terms of disciplinary origins and intellectual orientations of the authors.³ Among the perspectives which

fall within this group are theories of dual economy, dual labor markets, and labor force segmentation. Despite this diversity, there are certain common elements which characterize this approach and which differentiate it from neoclassical models. Primary among these is the proposal of a division of the industrial structure into distinct sectors (we will focus here on a two-sector model consisting of *core* and *periphery* sectors) within which employers and workers face fundamentally different conditions and operate according to fundamentally different rules.

Given this common emphasis on economic sectors there remain important variations in the definition and conceptualization of these sectors. Our approach is similar to that of Averitt (1968), Bluestone et al. (1973), and other writers in the dual economy tradition in that we conceptualize economic sectors as structural entities which derive from the nature of modern industrial capitalism (Beck and Horan, 1978). This usage differs from Piore (1975) and others in the dual labor market and labor force segmentation tradition who tend to define sectors, or segments, on the basis of the characteristics of labor markets and worker behavior (cf. Harrison, 1974; Spilerman, 1977). From our perspective, these labor market characteristics are seen as predictable outcomes of the sectoral structure, not as their defining characteristics.

In dual economy theory, the sectoral dichotomization is linked to the emergence during the late nineteenth and early twentieth centuries of a core industrial sector dominated by large corporate enterprises which came to constitute an oligopolistic system of production (Baran and Sweezy, 1966). The core sector thus is differentiated from the periphery sector which is characterized by smaller firms, operating in a more or less open, competitive capitalistic environment (Averitt, 1968). Bluestone et al. (1973:28–9) have summarized the characteristics of these two sectors of the economy as follows:

The core economy includes those industries that comprise the muscle of American economic and political power.... Entrenched in durable manufacturing, the construction trades and to a lesser extent, the extraction

³ See Cain (1976) for a review of the literature.

industries, the firms in the core economy are noted for high productivity, high profits, intensive utilization of capital, high incidence of monopoly elements, and a high degree of unionization. What follows normally from such characteristics are high wages. The automobile, steel, rubber, aluminum, aerospace, and petroleum industries are ranking members of this part of the economy. Workers who are able to secure employment in these industries are, in most cases assured of relatively high wages and better than average working conditions and fringe benefits....

Beyond the fringes of the core economy lies a set of industries that lack almost all of the advantages normally found in center firms. Concentrated in agriculture, nondurable manufacturing, retail trade, and sub-professional services, the peripheral industries are noted for their small firm size, labor intensity, low profits, low productivity, intensive product market competition, lack of unionization, and low wages. Unlike core sector industries, the periphery lacks the assets, size, and political power to take advantage of economies of scale or to spend large sums on research and development.

Theories of dual economy suggest that these sectoral differences have important implications for the opportunity structures and experiences faced by individual workers. In the core sector, workers move within job structures characterized by differentiated task and wage schedules with often well-defined career patterns, i.e., internal labor markets (Doeringer and Piore, 1971; Spilerman, 1977). Formal education is widely used to mediate individual access to job ladders, and workers' wages "are largely determined by their respective access to different job clusters, by the relatively rigid pattern of wages attached to the job structures through which they respectively move, and by the speed with which they pass through those structures" (Gordon, 1972:50). In the peripheral sector, occupational opportunity structure is more restricted with a consequent dampening of task and wage variations. Gordon (1972:51) suggests that in this sector, "variation in individual hourly wages will depend very little on variations in individual 'capacities' like aptitude, reasoning and vocational skill."

If this sectoral specification of the structure of the socioeconomic order is correct, it has important implications for the

analysis of individual earnings data. Specifically it implies that those analyses of earnings determination which assume a homogeneous market condition (i.e., the absence of sectoral differentiation) and attempt to estimate the earnings returns to individual characteristics such as schooling, social background, and work experience, will produce results which are systematically biased through misspecification of the economic structure. If returns are different in the two sectors, then no single estimate of, say, the return to schooling, will suffice (Beck and Horan, 1978).

In providing a nonindividualistic alternative for the analysis of earnings, theories of the dual economy provide structural explanations for economic differences between racial and sexual groups. Rather than interpreting group differences in earnings or poverty as due to different rates of individual "failure" in a competitive market, the theory suggests that such group differences may be the outcome of differential assignments of group members within the sectoral structure of the economic order (Gordon, 1972; Tussing, 1975; Bluestone et al., 1973; Edwards et. al., 1975; Harrison, 1974).

The analysis which follows is addressed to an empirical evaluation of a sectoral model of earnings determination. In undertaking this analysis we will be seeking empirical answers to three critical questions:

- (1) Are there differences in the labor force composition of core and periphery sectors?
- (2) Can differences in individual earnings between sectors be accounted for in terms of characteristics of the individuals located within those sectors?
- (3) How similar are the processes of earnings determination for core and periphery sectors?

An affirmative answer to the first question would be consistent with the findings of Osterman (1975) and Bibb and Form (1977) that worker characteristics such as income, education, race, and sex vary between economic sectors. Nonetheless, earnings differences could still be accounted for by differences in the "qual-

ity" of labor as reflected by variation in individual skills, social background, and abilities. Questions two and three constitute a departure from recent analyses of economic sectors in that they permit an evaluation of the meaning of sectoral differences for the basic earnings processes. If individual earnings discrepancies continue to exist after relevant labor force variables have been controlled, then the concept of economic sector must be acknowledged to be more than just a shorthand measure for differences in labor force quality. If substantial sectoral differences are found in the process of individual earnings attainment, this would call into question those status attainment parameter estimates which are based on the implicit assumption that sectoral differences are of no consequence for individual attainment. Such findings would provide strong evidence that factors relating to the organization of the industrial economy must be included along with individual-level variables in models of socioeconomic attainment.

DESCRIPTION OF DATA AND VARIABLES

The data used in this analysis are derived from the National Opinion Research Center's General Social Surveys for 1975 and 1976 (Davis, 1975; 1976). Included in the analysis are persons who reported working full or part-time, not working because of temporary illness, a strike or annual leave, and those who were laid off or unemployed and seeking work. Of these 1,695 respondents, information on industrial classification was unavailable for 12 and these cases were deleted leaving an n of 1,683. Excluded from the analysis are inexperienced unemployed and persons not in the labor force, such as housewives, students, and retirees. The use of this natural random sample of the experienced civilian labor force represents a departure from traditional stratification studies in several respects. Both males ($n=1,004$) and females ($n=679$) as well as whites ($n=1,512$) and nonwhites ($n=171$) are included. Unemployed persons with work experience ($n=120$) also are included in the analysis.

Researchers from "individualist" and

"structuralist" persuasions have typically differed in their choice of variables to include in models of earnings determination. Our analysis includes several variables which are characteristic of the individualist perspective. These include parental education⁴ and parental occupational prestige⁵ which may be viewed as measures of the background resources and socialization which a worker brings into the marketplace, years of age which may be interpreted as a proxy measure of labor market experience, and respondent's education which may be interpreted as an investment in future earnings. We have used two types of educational measures in the present study. In addition to the number of years of schooling completed, we have included a set of binary variables for the respondent's highest degree attained.⁶ These binary variables will be used to measure the extent to which levels of educational certification have an impact on earnings beyond that of years of schooling.

Another set of variables included in our analysis may be open to interpretational disputes. Sex and race could be included directly within the human capital framework by assuming that sex and race differences reflect basic inequalities in job-related capabilities. Few sociologists would make such an assumption, so some researchers have treated earnings differences between the sexes and between racial groups as due to market imperfections or to discrimination in the otherwise open, competitive system of earnings determination. Analysts in the structuralist

⁴ Parental education is taken as the years of schooling for the respondent's father. If this information is not available, the reported data on the mother's education is used. In the event that data are missing for both items, the mode of the distribution (eight years) is assigned.

⁵ Parental occupational prestige is measured using father's occupational data. Substitution of scores for missing data on this prestige variable is accomplished by random assignment of one of the three modes of the distribution. NORC occupational prestige scores are utilized.

⁶ There are four categories of degree attainment: (1) no high school diploma; (2) a high school diploma or junior college degree; (3) a bachelor's degree; and (4) a postgraduate degree. For the analysis presented here, the *no high school diploma* category is defined as the excluded category.

Table 1. Sectoral Classification of Industries

Core Sector	Periphery Sector
Mining	Agriculture, forestry, and fisheries
Construction	Durable manufacturing
Durable manufacturing	Lumber and wood products, except furniture
Stone, clay and glass products	Furniture and fixtures
Metal industries	Miscellaneous durable manufacturing
Machinery, except electrical	Nondurable manufacturing
Electrical machinery, equipment, and supplies	Food and kindred products
Transportation equipment	Tobacco manufactures
Professional and photographic equipment, and watches	Textile mill products
Ordnance	Apparel and other fabricated textile products
Nondurable manufacturing	Leather and leather products
Paper and allied products	Not specified nondurable manufacturing
Printing, publishing, and allied industries	Retail trade
Chemicals and allied products	Business and repair services
Petroleum and coal products	Personal services
Rubber and miscellaneous products	Entertainment and recreation services
Transportation	
Communications	
Utilities and sanitary services	
Wholesale trade	
Finance, insurance, and real estate	
Professional and related services	
Public administration	

tradition interpret the existence of sex and race differences in earnings as indicative of systematic forces involving differential opportunity structures which are embedded in the socioeconomic order.

A set of occupational variables which may be interpreted as characteristics of the market conditions encountered by an individual worker also are included in the model. These are occupational prestige, union membership, employment status, and work stability.⁷ The indicator of work stability—whether or not the respondent has been unemployed in the last ten years—is a weak measure of job continuity, but it is the only such measure available from the NORC data.

A crucial variable in our analysis is the distinction between core and periphery

economic sectors. The importance of this variable warrants a more detailed consideration of our core/periphery measure and a comparison of this with other sectoral classifications. Following Bluestone and colleagues (1973), as quoted above, we allocate to the core sector those industries which exhibit high levels of capital intensity, unionization, large assets, high profit margins, product diversification, and market concentration (see Table 1). These include mining, construction, some durable and nondurable manufacturing, transportation, communications, utilities, wholesale trade, finance, professional services, and public administration.

Industries were assigned to the periphery sector because of their small firm size, seasonal and other variations in product supply and demand, labor intensity, weak unionization, and low assets—all characteristics which Bluestone et al. (1973) attribute to competitive capitalism. The periphery industries include agriculture, portions of durable and nondurable manufacturing, retail trade, business and repair, and personal and entertainment services.

The sectoral dichotomy used in this paper is related to the distinction used by

⁷ It may seem unusual to have current unemployment status (employed or not employed) as an independent variable predicting the previous year's annual earnings since the respondent may or may not have been employed at that time. We are using current unemployment status, however, as an indicator of the vulnerability of the worker to changes in the demand for labor. The assumption is that, all other things equal, a worker who is currently unemployed is more likely to have been unemployed in the past. Cohn (1977) provides evidence of the reasonableness of this assumption.

Bibb and Form (1977) to classify blue-collar workers by economic sector. Our scheme is also related to Hodson's (1977) classification of industries into the three clusters suggested by O'Conner. O'Conner's (1973) neo-Marxist model of the economy posits that approximately one-third of the labor force can be found in each of three sectors: monopoly, competitive, and state.

Table A.1 in the Appendix contrasts the classificatory scheme used here with those of Bibb and Form and Hodson. While there are some differences in the sectoral placement of certain industries, there is consensus on the location of the majority. There are five instances in which our placement disagrees with both Bibb and Form and Hodson: printing, publishing, and allied industries; wholesale trade; finance, insurance, and real estate; professional and related services; and miscellaneous durable manufacturing. We allocated the first four of these industry clusters to the core sector, while Bibb and Form and Hodson placed them in the competitive periphery. We based this decision on tendencies toward oligopoly in the printing and publishing industries, the capital intensity of wholesale trade, the economic centrality of the finance industries as managers and brokers of capital wealth, and the strong control over the supply and prices of services exercised in the professions.

In the present analysis we treat the core/periphery distinction as being qualitative; that is, we assume that the two sectors are relatively internally homogeneous with respect to defining characteristics such as assets, market power, and the like. This is, of course, an oversimplification. However, if we can demonstrate the importance of this admittedly crude measure of differentiation in the economic structure, this will provide a rationale for work on a more refined measure of sectoral differentiation.⁸

The dependent variables in the analysis

⁸ We currently are involved in a project designed to refine the sectoral classification scheme. This research will lead to a more precise, empirically-grounded definition of the core and periphery sectors.

are the natural logarithm of the respondent's annual earnings and a binary variable coded one for earnings below the single person poverty threshold as defined by the U.S. Bureau of the Census (1976:146). Figures for 1974 annual earnings and the poverty threshold are adjusted into 1975 dollar units by the Consumer Price Index.⁹ Since earnings data were reported categorically, the midpoint of the response interval is used as a point estimate of the respondent's earnings. The midpoint for the highest, open-ended category (\$25,000 and over) was estimated by a Pareto approximation (Shryock and Siegel, 1975:366).¹⁰

ANALYSIS AND INTERPRETATION

To answer our first question—whether there are differences in the composition of core and periphery labor forces—the marginal characteristics of the two sectors are presented in Table 2. Core workers have larger, more homogeneous annual earnings than do their peripheral counterparts who have an average annual earnings disadvantage of \$3,057.97.¹¹ The degree of earnings inequality, as indexed by the Gini coefficient, is substantially larger in the periphery. Furthermore, the marginal probability of a peripheral worker having poverty earnings is more than twice that of the worker in the core, 0.3298 vs. 0.1564. Clearly, the economic position of workers in the peripheral sector is substantially inferior to that of workers in the core.

The human capital variables also exhibit substantial sectoral differences. Core workers, on the average, have more schooling, have better educational credentials, have parents who are better educated and with higher occupational status, and are more likely to be male and white than female and nonwhite. There is little

⁹ The single person, aged 14 to 64, poverty threshold for the 1974 earnings data in the 1975 NORC survey is \$2,557.00. For the 1975 earnings reported in the 1976 NORC survey, the threshold is adjusted to \$2,761.00 on the basis of the change in the Consumer Price Index.

¹⁰ Missing earnings data are assigned the modal value, \$12,500.

¹¹ $\text{Exp}(8.5993) - \text{exp}(7.7706) = 3057.97$.

Table 2. Means and Standard Deviations On Characteristics of 1,683 Workers in the Experienced Civilian Labor Force by Industrial Sector

Worker Characteristic	Industrial Sector		
	Core N=1125	Periphery N=558	t Ratio
Ln annual earnings	8.5993 (2.1148)	7.7706 (2.5628)	6.60**
Gini coefficient for earnings	0.397 (0.3632)	0.524 (0.4701)	—
Proportion poverty workers	0.3698 (0.4828)	0.4713 (0.4992)	-3.97**
Proportion nonwhite	0.0916 (0.2885)	0.1219 (0.3272)	-1.86*
Age for males ¹	39.63 (13.30)	38.81 (15.19)	0.81
Age for females ²	37.91 (12.92)	39.45 (14.45)	-1.41
Years schooling	12.85 (2.97)	11.44 (2.79)	9.55**
Proportion with less than a high school diploma	0.2133 (0.4096)	0.3620 (0.4806)	-6.27**
Proportion with a high school or junior college degree	0.5556 (0.4969)	0.5556 (0.4969)	0.00
Proportion with a bachelor's degree but no more	0.1582 (0.3649)	0.0753 (0.2639)	5.32**
Proportion with a post-graduate degree	0.0729 (0.2599)	0.0072 (0.0845)	7.70**
Proportion employed less than full-time	0.1129 (0.3165)	0.2634 (0.4405)	-7.20**
Proportion currently unemployed	0.0729 (0.2600)	0.0681 (0.2519)	0.36
Proportion unemployed at least once in past ten years	0.3307 (0.4705)	0.3423 (0.4745)	-0.47
Hours per week normally worked on current job ³	40.04 (11.65)	38.16 (16.05)	2.46**
Occupational prestige of current or last job	42.11 (14.10)	33.36 (11.64)	13.51**
Proportion belonging to a union	0.2969 (0.4569)	0.1380 (0.3449)	7.96**

Table 2. (Continued)

Worker Characteristic	Industrial Sector		
	Core N=1125	Periphery N=558	t Ratio
Parental years schooling	9.63 (3.93)	8.97 (3.70)	3.32**
Father's occupational prestige	39.65 (11.96)	38.41 (11.12)	2.10**

* Significant at 0.10 level.

** Significant at 0.05 level.

¹ N_c=709; N_p=295.² N_c=416; N_p=263.³ Only for those currently employed and working: N_c=1,004; N_p=505.

evidence, however, to suggest significant sectoral differences in average age for either males or females.

In terms of occupational characteristics, core workers are more likely than periphery workers to be in higher prestige occupations, to be employed full-time, to work more hours per week, and to belong to a union. It is interesting to note that those in the periphery not only work fewer hours, but also exhibit a greater dispersion in hours worked. This greater variance in the length of the average work week reflects both the larger proportion of part-time jobs in the periphery and a greater variation in the demand for labor in the periphery sector. There are no significant differences between sectors in either the current unemployment rate for the experienced labor force or in the proportion of workers who have been unemployed at least once in the past ten years, a weak indicator of work stability. This lack of difference may be due more to an insensitivity of the empirical measures than to true similarities in the unemployment experiences of core and periphery workers.¹²

¹² The unanticipated statistical equality of unemployment between sectors deserves some additional comment since one of the explicit hypotheses of the dual economy perspective is that there will be a higher degree of unemployment and underemployment in the periphery. If the proportion of less than full-time employment is taken as a measure of underemployment, the data in Table 2 show that the sectoral hypothesis is confirmed, yet there is no evidence that the rate of unemployment differs between

It is evident from these summary data that there are statistically and substantively important differences in labor force composition, work experiences, and earnings between the core and periphery sectors. This juxtaposition of intersectoral differences in earnings and poverty with significant differences in labor force composition and work experiences is suggestive of the importance of sectoral distinctions and justifies a further inquiry into the implications of industrial segmentation.

While the data in Table 2 clearly substantiate the differences between the marginal characteristics of the labor force in core and periphery, two questions remain to be answered. Can the sectoral discrepancy in annual earnings and the likelihood of poverty be accounted for by the differences in the "quality" of the two labor forces? That is, are the economic differences between sectors merely a function of differences in their respective labor forces, or are there persistent discrepancies which cannot be explained by labor force differences? If controlling on the labor force characteristics eliminates sectoral discrepancies in economic status, then we can conclude that such sectoral disparities can be attributed to labor force differences. On the other hand, if the economic discrepancies continue after adjusting for worker characteristics, we can infer that there are structural factors which influence economic status over and above the differences in the labor forces.

The third question which our analysis must address is that of differences in the processes by which economic outcomes are determined in the two sectors. To what degree do the mechanisms which affect economic well-being operate differently in the core and periphery? Such differences in basic economic processes would appear in our analysis as interactions between the sectoral variable and various worker characteristics.

sectors. Since there is a high degree of seasonal variation in unemployment and inasmuch as the NORC surveys are conducted over a rather large span of time, typically in March and April, we cannot have great confidence in the reliability of the unemployment rates as given in Table 2. Hence, we do not take the sectoral equality in unemployment as strong disconfirming evidence.

To address these issues we regress annual earnings, in logarithmic form, and a binary variable for poverty earnings on sex, race, human capital, and occupational variables and a complete set of sectoral interaction terms—defined as the product of each independent variable and a binary variable for industrial sector.¹³ The Chow test for the null hypothesis of no significant interactions produces F-ratios that are significant at the 0.01 level (see last line, Table 3).¹⁴ To permit further examination of the differential rates of return, the null hypothesis of no difference between sectors is tested using conventional t-tests on the interaction terms. Variables with interaction terms significant at or above the 0.10 level are presented in Table 3 as having only sectorally unique effects. When the null hypothesis of no interaction cannot be rejected, a common slope based on the pooled regression for all workers is presented in addition. Hence, the findings displayed in Table 3 provide the information necessary to answer our second and third questions.

There are two related strategies for answering our second question: can sectoral differences be accounted for by differential labor force "quality"? First we can ask how much does it cost on an annual basis for the average worker to be located on the industrial periphery? Or in other words, how much would the average periphery worker gain if that worker were located in the core sector? If the sectoral differences we noted above are due to differences in labor force quality, the average periphery worker should not be substantially better off in the core. On the other hand, if sectoral placement has an importance which transcends labor force quality, the average periphery worker would be substantially better off in the core.

Using this strategy, we can compute an estimate of the cost of being in the periphery first by obtaining a predicted income

¹³ For a more complex specification of an earnings function involving the human capital variables, see Mincer (1974:81-96).

¹⁴ See Hanushek and Jackson (1977:124-9) for a discussion of the Chow test and related procedures.

Table 3. Sectoral Regressions on Average Annual Earnings and the Probability of Poverty Earnings

Worker Characteristic	Ln Annual Earnings			Poverty Earnings		
	Common	Sector-Specific Core	Periphery	Common	Sector-Specific Core	Periphery
Intercept		7.4268**	5.2473**		0.3512**	0.7187**
Sex (1=female)		-0.8232**	-0.1439	0.1898**	0.1313**	0.2813**
Race (1=nonwhite)		-0.4693**	0.3230	-0.0183	0.0089	-0.0812
Age for males	0.0208**	0.0219**	0.0174*		-0.0024**	0.0004
Age for females		0.0185**	-0.0039	-0.0017	-0.0016	-0.0015
Years of schooling		-0.0635	0.1479**	0.0005	0.0110	-0.0149
Highest degree						
High school or junior college		0.4767**	0.3089		-0.1012**	-0.1030*
Bachelor's degree		1.0397**	-0.6999		-0.1519**	0.0534
Postgraduate degree		1.4708**	1.2753		-0.2186**	-0.1187
Occupational prestige	0.0246**	0.0217**	0.0266**		-0.0044**	-0.0081**
Union member (1=yes)	0.6490**	0.5343**	0.7725**		-0.1135**	-0.2195**
Unemployed (1=yes)		-0.6972**	-1.9885**	0.1351**	0.1363**	0.1723**
Work stability (1=stable)	0.4677**	0.3776**	0.6234**		-0.0699**	-0.1469**
Years parental schooling		-0.0109	-0.0802**		0.0028	0.0222**
Parental prestige	-0.0047	-0.0044	-0.0027	-0.0003	0.0006	-0.0022
R ² x 100		29.62			20.40	
F-ratio ¹		3.126			3.270	

*Significant at the 0.10 level.

** Significant at the 0.05 level.

¹ F-ratio for test of null hypothesis of no sectoral interactions, df=15, 1654.

based on applying the peripheral slopes in Table 3 to the means of the peripheral labor force, and then second by obtaining an expected income based on applying the core slopes in Table 3 to those same periphery means. This latter income is the expected annual earnings of the average periphery worker if that worker were in the core sector while the former is the income expected for the average periphery worker in the periphery. Subtracting these two expected incomes provides an estimate of the cost of being in the periphery for the average periphery worker. Using this procedure we find (see Table 4) that the periphery worker would gain

Table 4. Annual Costs and Benefits of Sectoral Location

Cost/Benefit	Dollars	Probability of Poverty Earnings
Cost of being in the periphery for the average peripheral worker	\$1,037.49	0.0890
Benefit of being in the core for the average core worker	\$979.46	0.0335

\$1,037.49 annually by being located in the core rather than in the periphery sector. Hence a change in sectoral placement without altering the racial, sexual, human capital, or occupational characteristics of the average periphery worker, would yield a substantial increase in annual earnings. Utilizing this same strategy but applying it to the poverty earnings equations, we find that the cost of being in the periphery (in terms of the probability of poverty earnings) is 0.0890. This means that if peripheral workers were in the core sector, they would have a poverty rate of 24.08% rather than their 32.98% rate (see Table 1). This 8.9 percentage point reduction indicates that the original sectoral discrepancy in poverty rates (17.34 percentage points) would be cut in half if peripheral workers received the same rates of return as their core sector counterparts.

Another strategy for answering our second question is closely related in that we ask how much benefit does the average core worker derive from being in the industrial core. Or alternatively, how much would the average core worker lose if that same worker were in the peripheral sector? Again, if the sectoral discrepan-

cies are due to differences in labor force quality, we should find that the average core worker obtains minimal benefit from being in the core sector.

To arrive at an estimate of this benefit, we first obtain an expected income based on evaluating the core slopes in Table 3 with respect to the means of the core labor force, and a second expected income derived by evaluating the periphery slopes with respect to those same core means. Subtracting these two expected incomes provides an estimate of the annual dollar benefit accruing to the average core worker for being in the core rather than in the periphery sector. Using this procedure, we find that the average worker's annual earnings are \$979.46 greater than they would be if a worker of the same quality was located in the periphery rather than in the industrial core. Applying this approach to the poverty earnings equations, we determine that one of the benefits of being in the core is a 0.0335 *reduction* in the probability of having poverty earnings relative to that expected if that same worker were in the periphery.

In sum, the findings noted in Table 4 indicate that there is a considerable cost borne by periphery workers over and above that which we can account for by the quality of that labor. Likewise, there is a substantial gain for workers in the core which cannot be explained by their labor force characteristics. Specifically, the data in Table 4 show that regardless of which sector is taken as the point of comparison, approximately one-third of the original \$3,057.97 sectoral earnings discrepancy is due to differences between sectors in the rates of return on worker characteristics while two-thirds of this gap is explainable by sectoral differences in labor force quality. These findings provide a clear answer to our second question: there are persistent sectoral differences in economic outcomes which cannot be explained by the racial, sexual, human capital, or occupational characteristics of their respective labor forces.

The answer to our third question—Are there sectoral differences in the processes by which economic outcomes are determined?—is clearly affirmative since the Chow test noted above is significant

for the annual earnings variable as well as for the poverty earnings variable. The findings presented in Table 3 indicate that the effects of some of the predetermining variables are sector-specific while others have a common effect in both sectors.¹⁵ This demonstrates that the effects of some variables in the model are contingent upon the industrial sector in which the worker is located and thus disconfirms the fixed rate of returns hypothesis.

A detailed comparison of the slope coefficients in Table 3 reveals some interesting and statistically significant sectoral interactions in the earnings determination process. In the core there are significant negative effects of being female and being nonwhite on annual earnings. There are, however, no apparent sex or race main effects in the periphery. We consider this to be of substantial theoretical import and will return to this point for further discussion.

In both sectors increased age for males is associated with increased annual earnings, reflecting the cumulative effects of additional labor market experience. The absence of a sector-age interaction for males means that for each additional year of age, earnings will increase by the same *proportion* in each sector, but this does not indicate that there are no *real dollar* differences in these age effects. Since core workers have higher average earnings, the real dollar benefits due to increments in age are greater for males in the core industries than for those in the periphery. To the degree to which age is a proxy for labor market experience for females, it would appear that increased experience has no payoff in the periphery in contrast to the significant positive relationship between age and earnings for females in the core. This minimal gain realized from the accumulation of work experience may re-

¹⁵ Since the dependent variable is expressed in logarithmic form, a partial slope represents the net proportionate change in annual earnings resulting from a unit increase in the predetermining variable. Since the average annual earnings in the core sector is greater than the average in the periphery, the real dollar increase due to a unit change in the predetermining variable will be larger for core industry workers even though some of the rates of return are common to both sectors.

sult from females being disproportionately assigned to the low-skill jobs which characterize the periphery. Such jobs have flat learning profiles such that there is little or no increase in worker productivity with increased job experience (Harrison, 1974).

Turning to education we see that schooling, net of degree levels, has an important positive effect on annual earnings in the periphery but no significant effect in the core. In contrast, the net effects of levels of degree attainment are significant and positive in the core but nonsignificant in the periphery. This suggests the hypothesis that earnings return to education in the core sector rests on the acquisition of a formal degree, whereas in the peripheral sector economic benefits are derived from additional years of schooling, not from increases in formal levels of certification.

An examination of the occupational variables indicates that higher occupational prestige, union membership, and stable work all make significant contributions to higher annual earnings. In each sector these variables operate in a similar manner in the sense that there are no sectoral differences in the *proportional* increase in earnings. In terms of real dollar earnings, however, the benefits of high occupational prestige, employment, work stability, and union membership are greater for those employed by core industry firms than for those working in the periphery. Finally, while it might be expected that the social background variables would have a direct influence on economic status, we find little evidence of such effects. There is a very small negative relationship between parental schooling and annual earnings in the periphery, but this effect does not exist in the core. Parental occupational prestige has no demonstrable direct influence in either sector.

While annual earnings is an important measure of overall economic well-being, the dual economy literature has placed special emphasis on the lower range of the earnings distribution (see e.g., Gordon, 1972; Tussing, 1975; Piore, 1977). Table 3 presents the effects of the individual and occupational variables on the probability that a worker will have annual earnings

below the poverty threshold.¹⁶ There are important differences between economic sectors in the effects of worker characteristics on poverty earnings.

In both sectors female workers have a greater probability of being poor than do male workers. Contrary to our findings in the earnings equation, there is no empirical evidence in either sector that nonwhite workers are any more likely to be poor than white workers. Of course, this only represents the absence of a *net* effect of race. Since marginally a nonwhite worker is 1.3 times more likely to be poor than a white worker, the lack of a significant race effect indicates that there are no race effects after controlling on sector, occupational, and human capital variables.¹⁷

The age effect for males has a significant negative impact on the probability of poverty earnings in the core, while it is nonsignificant in the periphery. There is no evidence that increased age has any measurable influence for females in either sector. Thus, additional years of age, and presumably, increased work experience reduce the likelihood of poverty only for males in the core sector. Core female workers and all periphery workers have a greater risk of earnings below the poverty threshold, but this risk is unaffected by age. In general we find that the schooling variable has no effect on poverty earnings in either sector. The degree certification variables, however, tend to reduce the likelihood of poverty for those employed in the core sector. In the periphery, only the high school/junior college degree variable has a significant effect on poverty, while years of education per se has no such effect.

The "costs" of occupying a low prestige occupation, of not belonging to a union, and of work instability are significantly higher in the periphery than in the core, although the effects are significant in

¹⁶ There are certain technical difficulties present when estimating a regression with a binary dependent variable. One of these is that the assumption of homoscedasticity is difficult, if not impossible, to meet. Therefore, the t-ratios and F-ratios resulting from the poverty earnings regression must receive a cautious interpretation.

¹⁷ $Pr(\text{Poor/Nonwhite}) = 0.2690$ and $Pr(\text{Poor/White}) = 0.2077$.

both. These differential effects on poverty earnings may reflect the absence of social mechanisms by which wage rates and work continuity are protected in the periphery. The fact that union membership is the single most important factor in reducing the likelihood of poverty earnings for a worker in the peripheral sector provides further support for this interpretation. In both sectors, being unemployed increases the probability of poverty by a significant degree.

These findings provide striking evidence for the importance of sectoral segmentation of the American economic system. Not only can sectoral differences in economic outcomes not be reduced to differential labor force quality, these sectoral differences serve to condition the impact of both human capital and occupational variables. Even more striking are the findings concerning the interactions between economic sector, race, and sex. Non-whites and women face significant earnings disadvantages in the core, but not in the periphery. That is, in the periphery sex and race groups face no disadvantage except for the sectoral disadvantage which they share with white males.

One response to these findings by human capital researchers, who generally restrict their analysis to employed white males, might be that our analysis has confounded exogenously-determined racial and sexual effects with sectoral differences. This would imply that our findings could not be replicated for full-time employed white males. To address this issue, we repeat the analysis above for white, full-time employed males ($N_{\text{core}} = 523$, $N_{\text{periphery}} = 193$).¹⁸

Computing the marginal characteristics for these white males, we find the earnings disadvantage for labor in the periphery to be a statistically significant \$5,155.81 which is substantially larger than the marginal disadvantage noted in Table 2 for all peripheral workers, \$3,057.97. Similarly we find that the currently full-time employed white male in the periphery is 3.18

Table 5. Annual Costs and Benefits of Sectoral Location For Full-Time Employed White Males

Cost/Benefit	Dollars	Probability of Poverty Earnings
Cost of being in the periphery for the average peripheral worker	\$4,097.51	0.0880
Benefit of being in the core for the average core worker	\$2,903.56	0.0413

times more likely to have poverty-level earnings than is his core counterpart. These two summary statistics indicate that the economic impact of the sectoral distinction is *more* pronounced among employed white males than among workers in general.

Replicating the regression analysis for this subset of workers, we applied the Chow test for the null hypothesis of no sectoral interaction to the regressions for the annual earnings and poverty earnings variables and in both instances we were able to reject the null hypothesis at the 0.01 level of significance.¹⁹ Hence, there is strong empirical evidence supporting the hypothesis that the process of earnings determination differs between sectors even when the analysis is restricted to full-time employed white males.

In order to determine whether sectoral differences can be attributed to variation in labor force quality, we estimate the costs and benefits of sectoral location for the average full-time employed white male in each sector. Using procedures analogous to those described in detail previously, we find (Table 5) that it costs the average white male in the periphery \$4,097.51 annually to be in the periphery. That is, if that same male were in the core we would anticipate an annual gain in earnings of that amount. For the average full-time employed white male in the industrial core, we estimate that sectoral placement yields an annual benefit of

¹⁸ Because of space limitations we have not presented either the marginal distributions or regression coefficients for the full-time employed white males. These are available from the first author.

¹⁹ The F-ratio for the earnings equation is 2.716 with 11 and 695 degrees of freedom. For the poverty earnings equation, $F = 3.977$ with 11 and 695 degrees of freedom.

\$2,903.56 over and above the earnings expected if that male had been located in the periphery. In addition to these dollar costs and benefits, we note that there are small costs and benefits in the probability of having poverty-level earnings associated with sectoral placement.

Comparing the costs and benefits of sectoral location for full-time employed white males (Table 5) with the costs and benefits for all workers (Table 4), we find that the dollar costs of being in the periphery sector and dollar benefits of being in the core sector are substantially greater for white males than for workers in general. This is consistent with the hypothesis that minorities are differentially channeled into the periphery where the rate of return on their already lower average stock of human capital is less. Those white males who find themselves in the economic periphery apparently are evaluated in a similar fashion to these minorities, hence the sectoral penalty paid by full-time employed white males is particularly heavy. This interpretation is consistent with the lack of race and sex main effects in the periphery which we noted in Table 3. If the peripheral white males had been employed in the core they would have gained from not only the higher rate of return on human capital and occupational variables, but also from the racial and sexual earnings discrimination found in that sector.

In sum, the findings reported in Table 5 are quite unambiguous in support of the position that even among full-time employed white males, sectoral placement makes a critical difference in economic well-being above and beyond the effects of differences in labor force quality. This persistent disparity must be attributed to the structural organization of the industrial economy.

CONCLUSIONS

Our analysis examines the existence and the importance of industrial sectors, as hypothesized by the dual economy literature, on the process of earnings determination. Since the status attainment and neoclassical income models rest on the assumption of labor market homogeneity, our efforts to test a theoretically-derived

model of nonhomogeneity are of direct relevance for those research traditions. Using a distinction between core and periphery industrial sectors derived from Bluestone et al., (1973) we address our analysis to three basic questions: (1) Are there significant differences in the economic status and composition of the core and periphery labor forces? (2) Can sectoral differences in economic status be reduced to differences in the quality of their respective labor forces? And, (3) to what degree are the economic returns to sex, race, human capital, and occupation-labor force variables the same in both sectors?

The analysis provides clear answers to all three questions. The core and periphery sectors *do* exhibit significant differences in both earnings levels and in labor force composition. The sectoral differentials in earnings *cannot* be explained away by differences in labor force quality. The relationships between earnings and human capital as well as occupation-labor force variables *do* differ significantly between core and periphery sectors. Whether considering the biologically-fixed attributes of race and sex, the human capital variables, or the occupation-labor force variables, we find that the real dollar returns on these worker characteristics are greater in core industries than in periphery industries.

The importance of these findings for the neoclassical research tradition in stratification should be clear. In face of strong empirical evidence contrary to the implicit tenet of labor market homogeneity, parameter estimates which are conditional on this assumption can no longer be treated as appropriate bases for the construction of sociological theory or social policy. Such simplistic models lead to a serious misspecification and misrepresentation of the social processes underlying individual earnings determination. In contrast, the notion of economic sectors appears to hold substantial promise as a theoretical concept and as an exemplar for a research program aimed at identifying structural aspects of the socioeconomic order.

We further contend that the sectoral model employed here constitutes an essential element in understanding the proc-

ess of discrimination against minority groups in that it includes not only individual factors but also the organization of the economic structure.²⁰ Specifically, despite the differential representation of females and nonwhites in the peripheral sector, there is no statistically significant evidence of earnings discrimination in that sector after controlling on the predetermining variables. In the core sector, however, there is evidence of significant adverse race and sex main effects on earnings even after controlling on human capital and occupational variables. This is not to suggest that the sectoral distinction is only relevant for minority group members. Quite the contrary, our analysis for the subsample of full-time employed white males indicates that the sectoral disadvantages in both earnings and the likelihood of poverty earnings are *higher* for this group than for the experienced civilian labor force as a whole. Hence, we find support for the dual economy perspective regardless of whether interest is focused on workers in general or is restricted to the primary labor force of full-time employed white males.

What does all of this mean? It means that we should be very suspicious of any attempts to build models of occupational or earnings processes in industrial society

which consist exclusively of individual-level variables. Our analysis suggests that the rules which govern the distribution of socioeconomic benefits to individual workers are not uniform across all sectors of the economy. Analysts in the human capital and status attainment tradition have tended to interpret income and status differences as due to the application of a fixed set of rates-of-return to different mixes of individual background characteristics, skills, and experience. We have shown that these rates-of-return are not fixed, and that one important determinant of their variability is a distinction between core and periphery sectors derived from theories of the structure of industrial capitalism. In our view these findings constitute a challenge to models of the social and economic order which underlie the human capital and status attainment research traditions. While we would be the first to acknowledge the primitive character of our dichotomous model of sectoral differentiation, we see the present analysis as a starting point for a type of research which will link individual socioeconomic behavior to models of industrial structure.

APPENDIX

Table A.1 presents the sectoral classification used in this paper along with those adopted by Bibb and Form (1977) and Hodson (1977). For a comparison of these schemes see the body of the text. Items designated with a question mark represent industries where the sectoral location was ambiguous.

Table A.1. Sectoral Classifications

Industry Group	Present Analysis	Bibb-Form (1977)	Hodson (1977)
Agriculture, forestry and fisheries	Periphery	Excluded	Periphery
Mining			
Metal mining	Core	Core	Core
Coal mining	Core	Core	Core
Crude petroleum and natural gas	Core	Core	Core
Nonmetallic mining and quarrying	Core	Core	Periphery
Construction	Core	Core	Core
Durable Manufacturing			
Lumber and wood products	Periphery	Core	Periphery
Furniture and fixtures	Periphery	Core	Periphery
Stone, clay, and glass products	Core	Core	Core
Metal industries	Core	Core	Core
Machinery, except electrical	Core	Core	Core
Electrical machinery, equipment, supplies	Core	Core	Core
Transportation equipment	Core	Core	Core

²⁰ In further research we are pursuing the implications of industrial segmentation for race and sex discrimination using data on a sample of the experienced civilian labor force from the 1976 Current Population Survey (March supplement).

Table A.1. (Continued)

Industry Group	Present Analysis	Bibb-Form (1977)	Hodson (1977)
Professional and photographic equipment	Core	Core	Core
Ordnance	Core	Core	Core
Miscellaneous durable manufacturing	Periphery	Core	Core
Nondurable Manufacturing			
Food and kindred products	Periphery	Periphery	Core
Tobacco manufacturers	Periphery	Periphery	Core
Textile mill products	Periphery	Periphery	Periphery
Apparel and other fabricated textiles	Periphery	Periphery	Periphery
Paper and allied products	Core	Periphery	Core
Printing, publishing and allied industries	Core	Periphery	Periphery
Chemicals and allied products	Core	Core	Core
Petroleum and coal products	Core	Core	Core
Rubber and miscellaneous plastic products	Core	Core	Core
Leather and leather products	Periphery	Periphery?	Periphery
Not specified nondurable manufacturing	Periphery	Periphery?	Periphery
Transportation			
Railroads and railway express service	Core	Core	Core
Street railways and bus lines	Core	Core	Periphery
Taxicab service	Core	Core	Periphery
Trucking service	Core	Core	Periphery
Warehousing and storage	Core	Core	Periphery
Water transportation	Core	Core	Periphery
Air transportation	Core	Core	Core
Petroleum and gasoline pipelines	Core	Core	Periphery
Services incidental to transportation	Core	Core	Periphery
Communications			
Radio broadcasting and television	Core	Core	Periphery
Telephone (wire and radio)	Core	Core	Core
Telegraph (wire and radio)	Core	Core	Core
Utilities and sanitary services			
Electric light and power	Core	Core	State
Gas, steam and supply systems	Core	Core	State
Electric-gas utilities	Core	Core	State
Water supply	Core	Core	Periphery
Sanitation services	Core	Core	Periphery
Other not specified utilities	Core	Core	Periphery
Wholesale trade	Core	Periphery	Periphery
Retail trade	Periphery	Periphery	Periphery
Finance, insurance, and real estate	Core	Periphery	Periphery
Business and repair services	Periphery	Periphery	Periphery
Personal services	Periphery	Periphery	Periphery
Entertainment and recreation services	Periphery	Periphery	Periphery
Professional and related services	Core	Periphery	Periphery
Public administration	Core	Core	State

REFERENCES

- Althausen, Robert P. and Arne L. Kalleberg
1977 "Occupational and firm labor markets: toward a specification of structural factors in wage attainment." Paper presented at the annual meeting of the American Sociological Association, Chicago.
- Averitt, Robert T.
1968 *The Dual Economy: The Dynamics of American Industry Structure*. New York: Horton.
- Baran, Paul A. and Paul M. Sweezy
1966 *Monopoly Capital*. New York: Monthly Review Press.
- Beck, E. M. and Patrick M. Horan
1978 "The structure of American capitalism and status attainment research: a reassessment of contemporary stratification theory." Unpublished paper, Department of Sociology, University of Georgia, Athens.
- Becker, Gary S.
1964 *Human Capital*. New York: National Bureau of Economic Research.
- Bibb, Robert and William H. Form
1977 "The effects of industrial, occupational,

- and sex stratification on wages in blue-collar markets." *Social Forces* 55:974-96.
- Bluestone, Barry, William M. Murphy, and Mary Stevenson
1973 *Low Wages and the Working Poor*. Ann Arbor: Institute of Labor and Industrial Relations, University of Michigan.
- Cain, Glen G.
1976 "The challenge of segmented labor market theories to orthodox theory." *Journal of Economic Literature* 14:1215-57.
- Cohn, Richard M.
1977 *The Consequences of Unemployment on the Self*. Ph.D. dissertation, Department of Sociology, University of Michigan. Ann Arbor.
- Davis, James A.
1975 *National Data Program for the Social Sciences: Codebook for the Spring 1975 General Social Survey*. Chicago: National Opinion Research Center.
1976 *National Data Program for the Social Sciences: Codebook for the Spring 1976 General Social Survey*. Chicago: National Opinion Research Center.
- Davis, Kingsley and Wilbert E. Moore
1945 "Some principles of stratification." *American Sociological Review* 10:242-9.
- Doeringer, Peter B. and Michael J. Piore
1971 *Internal Labor Markets and Manpower Analysis*. Lexington: Heath.
- Duncan, Otis D., David Featherman, and Beverly Duncan
1972 *Socioeconomic Background and Achievement*. New York: Seminar Press.
- Edwards, Richard C., Michael Reich, and David M. Gordon
1975 *Labor Market Segmentation*. Lexington: Heath.
- Gordon, David M.
1972 *Theories of Poverty Underemployment*. Lexington: Heath.
- Grandjean, Burke D.
1975 "An economic analysis of the Davis-Moore theory of stratification." *Social Forces* 53:543-52.
- Hanushek, Eric A. and John E. Jackson
1977 *Statistical Methods for Social Scientists*. New York: Academic Press.
- Harrison, Bennett
1974 "The theory of the dual economy." Pp. 269-87 in Bertram Silverman and Murray Yanowitch (eds.), *The Worker in "Post Industrial" Capitalism*. New York: Free Press.
- Hauser, Robert M. and David L. Featherman
1977 *The Process of Stratification: Trends and Analyses*. New York: Academic Press.
- Hodson, Randy D.
1977 *Labor Force Participation and Earnings in the Core, Peripheral, and State Sectors of Production*. M. A. thesis, Department of Sociology, University of Wisconsin, Madison.
- Horan, Patrick M.
1978 "Is status attainment research atheoretical?" *American Sociological Review* 43:334-41.
- Hudis, Paula M. and Arne L. Kalleberg
1977 "Labor market structure and sex differences in occupational careers." Paper presented at the annual meeting of the American Sociological Association, Chicago.
- Kalacheck, Edward and Frederick Raines
1976 "The structure of wage differentials among mature male workers." *Journal of Human Resources* 35:484-506.
- Mincer, Jacob
1974 *Schooling, Experience, and Earnings*. New York: National Bureau of Economic Research.
- O'Connor, James
1973 *The Fiscal Crisis of the State*. New York: St. Martin's.
- Osterman, Paul
1975 "An empirical study of labor market segmentation." *Industrial and Labor Relations Review* 28:508-23.
- Parsons, Talcott
1940 "An analytical approach to the theory of social stratification." *American Journal of Sociology* 45:841-62.
- Pease, John, William H. Form, and Joan H. Rytina
1970 "Ideological currents in American stratification literature." *American Sociologist* 5:127-37.
- Piore, Michael
1975 "Notes for a theory of Labor Market stratification." Pp. 125-50 in Richard C. Edwards, Michael Reich, and David M. Gordon (eds.), *Labor Market Segmentation*. Lexington: Heath.
1977 "The dual labor market: theory and implications." Pp. 93-7 in D. M. Gordon (ed.), *Problems in Political Economy: An Urban Perspective*. Lexington: Heath.
- Sewell, William H. and Robert M. Hauser
1975 *Education, Occupation and Earnings: Achievement in the Early Career*. New York: Academic Press.
- Shryock, Henry and Jacob S. Siegel
1975 *The Methods and Materials of Demography*. Washington, D.C.: U.S. Government Printing Office.
- Spilerman, Seymour
1977 "Careers labor market structure, and socioeconomic achievement." *American Journal of Sociology* 93:551-93.
- Stolzenberg, Ross M.
1975 "Occupations, labor markets and the process of wage attainment." *American Sociological Review* 40:645-65.
- Thurow, Lester C.
1975 *Generating Inequality*. New York: Basic Books.
- Tussing, A. Dale
1975 *Poverty in a Dual Economy*. New York: St. Martin's.
- U.S. Bureau of the Census
1976 "Characteristics of the population below the poverty level: 1974." *Current Population Reports, Series P60, No. 102*. Washington, D.C.: U.S. Government Printing Office.

POWER, EQUITY AND COMMITMENT IN EXCHANGE NETWORKS*

KAREN S. COOK
AND RICHARD M. EMERSON
University of Washington

American Sociological Review 1978, Vol. 43 (October):721-739

Exchange theory has the virtue of bringing both power and equity together in a single analytic framework. However, exchange theory has focused largely upon analysis of the dyad, while power and justice are fundamentally social structural phenomena. First, we contrast economic with sociological analysis of dyadic exchange. We conclude that (a) power and equity from social exchange theory carry us beyond economic theory of dyadic exchange; yet (b) for power and equity to be studied effectively, analysis of systems larger than the dyad is needed. Second, we introduce exchange networks to extend power and equity analysis into more macroscopic n-person social structures. Third, a laboratory method is reported for controlled study of exchange networks as bargaining structures. Finally, we present findings which show that (a) power is an attribute of position in a network structure observable in the occupant's behavior, even though the occupant does not know what position or what amount of power s/he possesses; (b) equity or justice concerns constrain the use of that power; (c) emergent interpersonal commitments impede the use of power; and (d) when power is unequally distributed among actors in a network, females form stronger commitments to their exchange partners than do males. In conclusion, we discuss the importance of commitment in distinguishing between economic and social exchange theory.

INTRODUCTION

We report here the first experiment in a series dealing with *social power* and *equity* in exchange networks. Our major objective in this experiment was to investigate the *social structural* determinants of power and *normative constraints* on the use of power. The research was conducted within an exchange framework because both power and equity have received major attention within social exchange theory.¹ For our purpose, ex-

change theory has the advantage of bringing power, as the capacity to exploit, and justice, equity or other normative restraints upon exploitation together in a single analytic framework.²

Past theory and research using the exchange approach have dealt with power at a distinctly microanalytic level. This is a theoretical flaw if social power is, as we believe it to be, a social structural phenomenon. In order to develop this point, we, beginning with the analysis of dyadic exchange, will compare economic with social conceptions. We then will extend the discussion to exchange and power in larger *network* structures, before presenting experimental results on the use of power in exchange networks.

THE ECONOMICS OF DYADIC EXCHANGE: INDETERMINANCY IN BILATERAL MONOPOLY³

Consider two persons engaged in trade with each other. Let Ax,By be an ex-

attention in exchange theory (Homans, 1961; Adams, 1965; Blau, 1964b).

² Recently a variety of researchers have suggested that relatively little attention has been given to the analysis of the linkage between power and equity or justice processes (see Lerner, 1975; Walster and Walster, 1975; and Homans, 1976).

³ There is if Economics a well developed science

*Address all communications to: Karen S. Cook; Department of Sociology; University of Washington; Seattle, WA 98195.

This research was supported by a grant from the National Science Foundation (SOC 75-04059). We gratefully acknowledge the research assistance of Mary Gillmore, Daniel Klepinger, Alan Mullendore and Barbara Sorensen. Their insightful critiques throughout the design and conduct of this research were especially valuable. The authors assume joint responsibility for the work reported here; the names have been listed alphabetically.

¹ Thibaut and Kelley (1959) introduced the concepts *behavior control* and *fate control*, and formulated power in terms of mutual dependence of exchange partners. Similarly, Homans (1961; 1974) and Emerson (1972) give central importance to power in their treatments of exchange theory. Blau (1964a) included power in the title of his major work on exchange; he argued that "exchange processes give rise to differentiation of power" (1964a:22). Principles of equity or justice also have received major

change relation between actors A and B who exchange resources of type x and y through trades involving variable quantities represented as X and Y . Economists often portray this two party exchange relation using the *edgeworth Box* (see Figure 1). If each party currently has some proportion of the total x and some proportion of the total y , then the pair is located at a point L prior to any trade. Two *indifference curves*, a_1a_3 for actor A and b_1b_3 for actor B, pass through L . These lines map the set of possible trades to which each actor is indifferent. Depending upon the values of x and y to A and B, these two curves might intersect as shown in Figure 1, in which case they enclose a zone containing all mutually beneficial trading possibilities (see shaded area in Figure 1).

Most important for our purpose is the line a_2b_2 in Figure 1, sometimes called the *contract curve*. Economic theory dictates that there be such a line defining maximum joint benefit.⁴ Economic theory also shows that exchanges will tend to occur somewhere on that line. The two traders A and B will explore potential transactions through offers and counter offers of amounts of x and y . Starting at point L , any offer within the shaded area in Figure 1 will be mutually beneficial and therefore potentially acceptable. However, any offer not on the contract curve can be improved for both parties by moving onto the line of maximum joint benefit. Thus, economists dealing with bilateral monopoly argue that a *bargaining path* will lead from L to some point E where exchange takes place somewhere along the contract curve. But in conventional microeconomic theory the course followed in bargaining and the eventual location of E along the line a_2b_2 are said to be "indeterminate."

We present these conceptions from microeconomics for two reasons. First, in

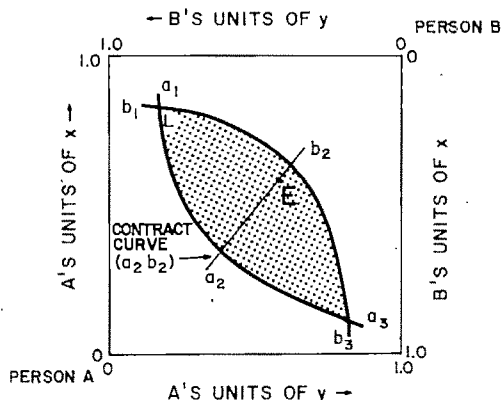


Figure 1. A Schematic Representation of Dyadic Exchange as Intersecting Indifference Curves, a_1, a_3 and b_1, b_3 , for Persons A and B

the experiment reported below we operationalize the contract curve and measure exchange along that line. Furthermore, manipulating the length of the contract curve, we create exchange relations with high and low potential benefit. The second reason for presenting the above conceptions is theoretical. Conventional microeconomic theory leaves two-person interaction indeterminate because it is a theory organized around *individual decision making* in an asocial market composed of many traders. But when only two parties engage one another in reciprocally tactical interaction, be they husband and wife or organized labor and management, the market mechanism is replaced by *social interaction* and economic theory is left with little to say about the outcome.⁵

THE SOCIOLOGY OF DYADIC EXCHANGE: POWER AND EQUITY

Social exchange theory has focused on the very phenomenon which economic exchange theory treats as indeterminate. Both power and equity, the two topics most central to social exchange theory, can be seen as dealing directly with the location of exchange outcomes along the contract curve a_2b_2 in Figure 1. That being the case, what can the sociology of equity and power contribute to a solution to the

of exchange. Therefore, if "social" exchange theorists are going to build another science of exchange, it is incumbent upon them to relate their work to economic theory, attending to and highlighting important differences.

⁴ The length of the line a_2b_2 can be taken to represent the amount of benefit attainable through exchange. It is properly taken as a *variable* along which exchange relations can be compared.

⁵ For a good summary of the history and present status of theoretical indeterminacy in bilateral monopoly see Coddington (1968).

economic problem of indeterminacy in dyadic exchange?

The first point of interest is that equity theory applies only when both A and B know, or believe they know, how much the other party benefits from a given transaction. Let αx and αy be the unit value (i.e., "preference for" or "utility") of the resources x and y to A. Similarly, let βx and βy be the unit value of resources x and y to B. Let X and Y refer to the number of units exchanged. In two-party exchange, equity can be specified quite simply in terms of equal profit:

$$\alpha x Y - \alpha x X = \beta x X - \beta y Y. \quad (1)$$

We refer to (1) as the rule of *equitable exchange*.⁶ Equity and distributive justice have been expressed differently by different researchers,⁷ but regardless of which equity formula is used, our main point is that equity theory offers a solution to the problem of indeterminacy. However, the solution requires that the following conditions are met:

Both A and B know or hold strong hypotheses about relative profit.

The rule of equitable exchange is activated and honored by A and B in the exchange situation at hand.⁸

⁶ This is only one of a family of distribution rules which emerge under different social conditions (see Cook and Parcel, 1977).

⁷ Homans (1961; 1974), Adams (1965) and Walster et al. (1973) offer proportional formulations of equity concerning the relationship between the inputs and outcomes involved in an exchange. In our opinion, proportional conceptions of equity or justice are more applicable to situations in which some central allocator or third party is distributing rewards among contributors than they are to direct two-party exchange. In a dyadic exchange in which two actors A and B each possess a single resource (x and y) and the exchange involves a trade of some quantity of x for some quantity of y , what one party inputs to the exchange is the other party's outcome. Thus a conception of equity in this case involves the relative comparison of the values of x and y to each actor and the quantities exchanged as specified in what we refer to as the *rule of equitable exchange*.

⁸ The value of x and y to actors A and B is conceived for our purposes as a location in a preference order (see Emerson, 1967a). The first condition is most likely to hold when (a) A and B share a common culture and (b) interact with one another repeatedly. The second condition is most likely to hold when A and B interact (a) with future interaction anticipated, and (b) when the stakes are high (i.e., when x and y are highly valued). Anthropological evidence on the last two points can be found in Sahlins (1965).

Given these conditions any proposed trade to which the rule of equitable exchange (1) does not apply is likely to be rejected. As Homans (1961; 1974) might put it, the proposed trade would make one party feel angry and leave the other party feeling guilty.

But suppose that the two conditions above do not hold. Assume, instead, that the subjects, having no knowledge of the other's benefit or no concern about equity or fairness, act on the basis of a "rational" distribution rule or exchange rule of maximum self-gain (Meeker, 1971). Under this condition social interaction will involve primarily a power process. Does power theory also offer a solution to the problem of indeterminacy? *Social power* in the dyad can best be conceived as a contest played out along the contract curve, a_2b_2 in Figure 1, with A pressing for outcomes in the direction of b_2 and B pressing toward a_2 . Let us define:

Def. 1. (Power): in any exchange relation $Ax;By$, the *power* of A over B (P_{ab}) is the potential of A to obtain favorable Y minus X outcomes at B's expense.⁹

Def. 2. (Dependence): the *dependence* of A upon B (D_{ab}) is a joint function (1) varying directly with the value of αy minus αx , and (2) varying inversely with the availability of y to A from alternative sources.¹⁰

⁹ There has been some discussion in the literature on power concerning problems introduced by a concept defined as a potential activity (cf. Bierstedt, 1950; Dahl, 1957; Wrong, 1968; Nagel, 1968). It is essential to a structural approach that power be defined as a potential to exert influence. Only then can we treat power as an attribute of *position* available for "use" by occupants and thus separate power from *power-use*, each subject to its own set of determining principles.

¹⁰ This definition warrants extensive discussion on several points we will only allude to here. First, while introduced here as a definition, it can be converted into two propositions and a joint function, all three to be advanced for empirical test. Second, the concept of value entails a preference order among qualitatively different resources. Third, and most important, the notion of *availability* will be treated below (a) in terms of simple access to alternative sources of y ; and (b) in terms of comparison levels (Thibaut and Kelley's "CL," 1959) based on a history of exchange with each source.

These definitions are taken, with minor modifications, from Emerson (1962), where they are linked in the theorem: $P_{ab} = D_{ba}$.

In power-dependence terms, a bilateral monopoly is the special case in which neither party has any alternative sources. Hence, relative power is based entirely upon the values αy and αx relative to βx and βy . If we assume that:

Both A and B use their power,

then it follows that:

X and Y will change over repeated transactions until equilibrium is reached at:

$$\alpha y Y - \alpha x X = \beta x X - \beta y Y. \quad (2)$$

This reasoning follows because the use of power to gain improved outcomes makes the user more dependent and therefore less powerful.

Thus, in bilateral monopoly, equity concerns and power processes make the same prediction concerning the outcome of the exchange (compare assertion (2) with the rule of equitable exchange (1) stated above). Whether through normative concerns about equity (given knowledge of other's benefit) or through selfish bargaining (blind to or unconcerned about benefit to the other party), the exchange outcome in a bilateral monopoly will tend to be the same. This is implied in the well known principle of *least interest* in bargaining power (Kuhn, 1964). The party who is receiving the least comparative benefit from a trade has the greater bargaining power to improve upon that trade. If that power is used, as we assume above, then the terms of the trade will shift until power is balanced (2); and in bilateral monopoly it is balanced at the point of equitable exchange (1).

Conclusions from Dyadic Exchange Theory

The theoretical considerations presented above lead to two conclusions. First, social exchange theory appears to solve the problem left indeterminate in economic theory concerning the outcomes in dyadic exchange or bilateral monopoly. The solution, however, will be useful only

when the theory can specify (a) the conditions under which people invoke and abide by a norm of equity, and (b) the conditions under which people use the (potential) power they have. In this experiment, therefore, we take *power use* as the main dependent variable and examine the effect of equity concerns upon an actor's use of his/her potential power.

Our second conclusion from dyadic exchange theory is that *power* and *equity* cannot be definitively studied in the dyad. Both concepts lose their fundamental social meaning if their outcomes converge as we have shown in the case of isolated two-party exchange. Power and normative restraints on power, such as equity, are fundamentally social structural phenomena properly studied in systems larger than the isolated dyad.¹¹

POWER, EQUITY AND COMMITMENT IN EXCHANGE NETWORKS

There are two different but complementary ways to extend exchange conceptions beyond the dyad to larger social systems. One involves corporate groups having two or more members occupying roles in a consensually defined role system. The other approach involves social structures composed of three or more actors located in positions within a connected network. The experiment reported here deals with exchange networks. Space will not allow a review of the rapidly growing literature on social networks, and presentation of a theory of network structure will be deferred. Rather, we provide only a few guiding definitions and comments required to introduce the experiment presented below.

It is fairly common in some network research (e.g., sociometric studies of affect structures and communication net-

¹¹ As Blau (1964:32) has concluded, "... the distinctive processes of power cannot be manifest in a dyad." This conclusion must not be taken to mean that the dynamics of bilateral monopoly are unimportant. For example, most collective bargaining situations entail features of bilateral monopoly (Pen, 1959). But such real-world cases of bilateral monopoly also entail third parties in roles which make network-structural analysis relevant (e.g., as mediators and as bargaining agents).

works) tacitly to assume that if two relations A-B and A-C share a member A in common then the relations can be joined to form the larger single structure B-A-C. This should not be taken as an assumption. For example, if A and B communicate, and A and C communicate, whether or not information gets from B to C through A remains to be demonstrated. Similarly, if A helps B and B helps C, whether or not C is (indirectly) helped by A remains to be demonstrated. Instead of making assumptions a priori, systematic study should be conducted concerning a class of empirical phenomena we will call *network connections*. Toward that end we define:

Def. 3. An *exchange network* is a set of two or more connected exchange relations.

Def. 4. Two exchange relations are *connected* to the degree that exchange in one relation is contingent upon exchange (or nonexchange) in the other relation.

(1) The connection is *positive* if exchange in one is contingent upon exchange in the other.

(2) The connection is *negative* if exchange in one is contingent upon nonexchange in the other.

Examples of positively connected networks range from Malinowski's (1922) Kula Ring to vertically integrated markets and channels of distribution (El-Ansary, 1972). Family lines can be analyzed similarly with kin-based economic opportunities flowing "down" and filial loyalty and support flowing "up" across generations. Examples of negatively connected networks include competitive economic market structures, dating networks and friendship networks. Most concrete networks involve connections of both types.¹² In this experiment we explore

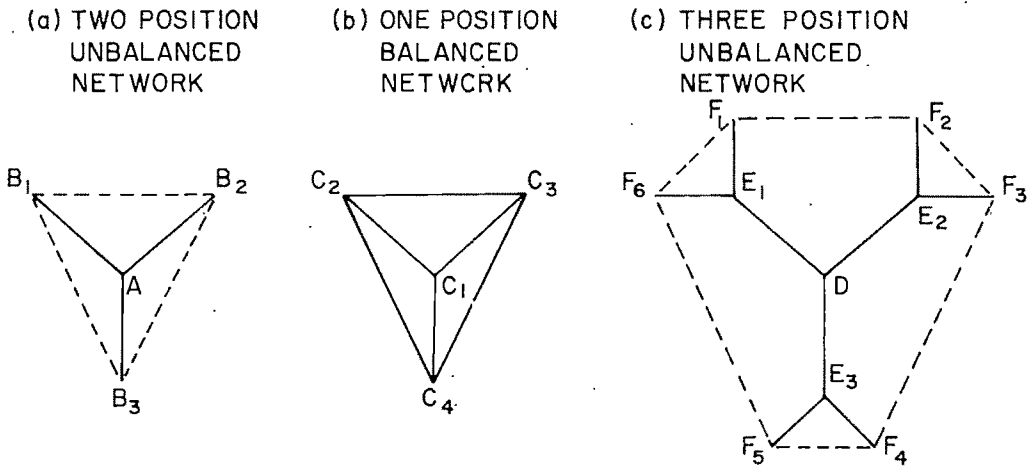
power use in networks based solely upon negative connections. We begin our studies with negatively connected networks because (a) the relation of power and dependence to position in negatively connected networks is relatively straightforward, and (b) negative connections are easily operationalized in the laboratory.

In the experiments underway in this series we manipulate network structures establishing power differentials in the process. Power use and normative restraints on power use are then studied over time as exchange takes place within the network. Figure 2 presents a graphic representation of the network structures under study. Each line represents an exchange relation between two persons. The exchange relations are negatively connected because we establish in the laboratory that all relations involving the same person are *alternatives* for that person (as referred to in Def. 2 above). Exchange with one partner precludes exchange with any other during a given time period. Since the network is a set of connected dyads, the concepts portrayed in Figure 1 apply to each line in Figure 2. Thus, those lines can be conceived as contract curves along which power can be exercised in a bargaining process. Persons joined by solid lines have 24 units of profit to divide through negotiations. Persons joined by broken lines have little value to exchange (eight units). As a result, as offers and counteroffers flow within the network a strong preference soon forms for exchange along the solid lines. Finally, all persons who have structurally identical locations within the overall structure are said to occupy the same *position* in the network.¹³ Positions are identified by let-

We then can describe complex networks based upon combinations of both positive and negative connections of varying strength. The concept of *network connection* allows us to envision a theory in which events happening at any location in such networks will have mathematically predictable and empirically traceable repercussions to all boundaries of that network. Clearly, such network connections must be taken as topics for empirical research in their own right.

¹³ People in structurally similar locations are said to occupy the same *position*. We will provide no explicit definition of position until the theory becomes more

¹² In addition to the sign of the contingency linking exchange relations, network connections can vary in strength. As a result networks have empirically definite boundaries at points where the contingency in Definition 4 is zero or, relatively speaking, very low.



KEY: All lines are communication links;
 Solid lines have high exchange value (total benefit=24)
 Broken lines have low exchange value (total benefit=8)
 Letters identify "positions" and numerical subscripts
 identify persons as occupants of positions.

Figure 2. Graphic Representation of Balanced and Unbalanced Network Structures (2a) Contains Two Positions, Four Persons (2b) Contains One Position, Four Persons (2c) Contains Three Positions, Ten Persons

ter and persons, as occupants of positions, are signified by numerical subscripts.

Definitions 1 and 2 regarding power and dependence now can be applied directly to positions. One of the variables governing dependence (*alternatives*) is represented in the number of high-value exchange relations the position provides. The other variable is held constant by controlling resource value so that, within each dyad, $\alpha y = \beta x$ and $\alpha x = \beta y$. That is, the two partners in any exchange relation have equal potential gain and equal potential loss. Therefore, applying Def. 2, relative dependence and therefore relative power as well are based entirely upon position in the network structure. The following assertions concerning power and position in the structures portrayed in Figure 2 can be derived from Def. 2 and the theorem $P_{ab} = D_{ba}$:

Power is balanced in relations between occupants of the same position:

formal, at which point it will be given a graph-theoretic definition. For present purposes it is sufficient to note that the residual graphs obtained by removal of persons holding the same position are identical graphs.

$P_{b_1b_j} = P_{b_jb_1}$; $P_{c_1c_j} = P_{c_jc_1}$; $P_{f_1f_j} = P_{f_jf_1}$; and power is unbalanced in relations between occupants of different positions. Specifically:

$$\begin{aligned} P_{ab_1} &> P_{b_1a}; \\ P_{e_1f_j} &> P_{f_je_1}; \text{ and} \\ P_{e_1d} &> P_{de_1}. \end{aligned}$$

These power inequalities are interesting for two reasons. First, the subjects located in the structure need not know either their position or their relative power for that power to operate in the ensuing exchange process. In the experiments in this series the subjects *have no knowledge about structural arrangements remote from their own location*. Therefore, positions A, C, D, and E are psychologically identical to each other, as are B and F; yet they provide their occupants with different levels of *structurally determined* power. The experiment measures in exchange behavior the *use* subjects make of the power their positions provide.

The second point of interest is specific to the assertion above that $P_{e_1d} > P_{de_1}$. This assertion is an unequivocal reversal of the notion that power is a function of the cen-

trality of one's position in a social system (cf. Hopkins, 1964). In Figure 2c, position D is the most "central" by current graph-theoretic measures (e.g., Freeman, 1977), yet by power-dependence principles, power will concentrate in position E in negatively connected networks.

Power Use in Exchange Networks

The exchange process taking place over time within a network structure will generate a reward distribution measurable as historically emergent Comparison Levels (CL_{ij}).¹⁴ For every line in Figure 2 and value of CL_{ij} and CL_{ji} will form over continuing exchange such that $0 \leq CL_{ij} \leq 24$ (or 8). Referring back to Def. 1 concerning power, it is clear that CL_{ij} is a direct measure of *power used* by person I over person J. Thus, our first task is to predict the emerging CLs throughout the network assuming total power use.

Let us specify the assumption of power use further by stating that power is used in the following way:

All members of the network *explore* alternative sources of benefit in the network (a) through extending offers to others and (b) by comparing offers and counteroffers from others.

All members of the network maximize benefit by (a) accepting the better of any two offers, (b) lowering offers when offers go unaccepted, and (c) holding out for better offers when it is possible to do so.

It is clear that two occupants of the same position have equal exchange opportunities. If their exchange with each other results in unequal CLs, and if the two assumptions listed above hold, then it follows that the party with the *lower* CL is the less dependent and the more powerful person in subsequent exchange between those two parties.¹⁵ Hence, in exchange

among actors in the same position in Figure 2:

Hyp. 1. CLs will change over successive transactions until equilibrium is reached at:

$$\begin{aligned} CL_{b_1b_j} &= CL_{b_jb_1} = 4; \\ CL_{c_1c_j} &= CL_{c_jc_1} = 12; \\ CL_{f_1f_j} &= CL_{f_jf_1} = 4. \end{aligned}$$

By the same reasoning we can now specify the equilibrium condition in the power unbalanced relations. If $CL_{ab_1} < 20$, then A will be less dependent and more powerful than B₁ in subsequent exchange (because some B₁ will offer more than CL_{ab_1}). Hence, in exchange among actors in the unbalanced structures in Figures 2a and 2c:

Hyp. 2. CLs will change over successive transactions until equilibrium is reached at:

$$\begin{aligned} CL_{ab_1} &= CL_{e_1f_j} = CL_{e_1d} = 20, \\ \text{and} \\ CL_{b_1a} &= CL_{f_je_1} = CL_{de_1} = 4. \end{aligned}$$

Stated more generally, equilibrium is reached for all structures in Figure 2 when for any three actors I, J and K, $CL_{ij} = CL_{ik}$. At this point no person has a choice among different offers. Persons in the powerful positions (A and E) receive only offers of 20; persons in weak positions (B, D, and F) receive only offers of four; and occupants of the balanced structure receive only offers of 12.

Normative Constraints on the Use of Power

Assuming that power is fully used, we have derived specific predictions concerning the outcome of bargaining in network structures. These predictions can be tested experimentally while at the same time using them as standards against which to measure the effect of normative or other constraints on the exercise of power.

1. *Equity*. In each of the exchange relations in Figure 2 resources will be given equal value. Hence, equity should obtain

¹⁴ We use CL_{ij} exactly as Thibaut and Kelley (1959) used CL. It refers to the level of Y recently received by I from J and therefore expected (in the predictive sense) by I in current exchange.

¹⁵ In applying the definition of dependence to network positions *availability* was taken to mean *potential* benefit from potential exchange partners. In the analysis of ongoing exchange, availability can be

measured in terms of actual benefit recently obtained (CL) or currently offered.

when $CL_{ij} = CL_{ji}$. Power use, however, will move exchange toward gross inequality in unbalanced relations generating pronounced inequity. Thus, (a) if *knowledge* of other's outcomes is provided in one experimental condition; and (b) if significant *equity concerns* are operating, then we predict:

Hyp. 3. Power use will be lower in the knowledge condition in unbalanced exchange relations than in the no-knowledge condition.

2. *Commitment*. The experiment reported below also was designed to explore a variable which we believe to be central in distinguishing social from economic exchange theory. The latter, being bound closely to the concepts *rationality* and the *perfectly competitive* market, carries the implicit assumption that exchange partners develop no loyalties or longitudinal commitments to one another. Insofar as these factors operate in reality, the real market structure is imperfect and the elegance of conventional economic theory is compromised in trying to deal with it. By contrast, reinforcement psychology and much of sociology and social anthropology tend to take the existence of longitudinal commitments as theoretically expected. Indeed, the perfectly competitive market of uncommitted actors must be seen sociologically, historically, and anthropologically as the deviant condition.¹⁶ Thus an additional aim of our research was to examine the possible role of emerging commitments among the subjects in the exchange networks as a determinant of level of power use within an exchange relation.

An actor is said to be committed to another actor in the network to the extent that choice of current exchange partner, from among alternative partners, can be predicted from previous partnerships. To the extent that commitments form, the exploration of alternatives is curtailed.

¹⁶ For discussion of this topic see Firth (1967:4) and Sahlins (1965:139) along with the entire thesis of Polanyi (1957).

Therefore it follows:

Hyp. 4. Power use varies inversely with commitment.

At the attitudinal level, commitment would involve irrationality in the short-run sense of ignoring better alternatives in favor of staying with old partners. However, in an uncertain or risky exchange environment, commitment might carry its own long-range utility.¹⁷

3. *Sex Differences in Power-Use*. We advance no theory of sex differences in the use of power. However, the literature contains many findings and discussions suggesting women might be (1) more concerned about equity, (2) more prone to form committed dyadic relations (see Eder and Hallinan, 1978), and (3) less likely to bargain in an assertive manner (see Rubin and Brown, 1975:169-74). All three of these hypotheses suggest lower power use by females. Where socialized normative restraints in the use of power blend off into sex role stereotypes in these discussions is impossible to say. However, in this research setting we have an ideal situation in which to observe whatever sex differences might exist.

EXPERIMENTAL PROCEDURES

Laboratory Apparatus

A newly developed computerized laboratory was employed for this study. It consists of a system of eight isolation booths containing CRT computer terminals linked together through a PRIME 300 minicomputer (for details see Cook and Emerson, 1977). Under programmed control, subjects may send messages to other accessible subjects. The computer serves both to relay and to record the messages. Subjects in this study were assured that no deception was used and that they were not dealing with the computer, but with one another through the computer.

¹⁷ There have been few discussions of the concept of *commitment* as used here (see Abramson, 1958; Huesmann and Levinger, 1976; Leik and Leik, 1972; Rubin and Brown 1973; and Marks, 1977). To our knowledge, this is the first experimental study dealing with this phenomenon.

Subjects

A total of 112 subjects (56 males and 56 females), recruited on the basis of desire to earn money, took part in a study of negotiated trade agreements. Eight subjects of the same sex met together with the experimenter for a brief orientation, followed immediately by a training session at the computer terminals. Two days later subjects returned to the laboratory for a three-hour experimental session. During that session, each of eight subjects was connected to three others as bargaining partners, forming two separate four-person networks. These *four-person networks are treated as the primary statistical unit of analysis throughout this study*.¹⁸

Overview of the Basic Design

The design was factorial-type containing two balanced between-subjects variables (sex and power) and two within-subjects variables (trials and equity). In addition, trials were nested within equity. Within levels of sex, subjects were randomly assigned to networks of four persons each. The groups were then randomly assigned to the power conditions (balanced and unbalanced).

Orientation and Training of Subjects

During the training session eight same-sex subjects met together for general instructions. In order to decrease the possibility of interaction outside the laboratory, subjects were asked during the orientation session whether they knew anyone in the group. If acquaintanceships or friendships existed, subjects were told that they would be assigned to different four-person groups during the experiment. They were told they would be paid \$3/hour during training, but that when they returned for the experimental session, they would be paid exactly what they earned by completing trade agreements

during the experiment. Subjects could not interact with each other except through their computer terminals. They were asked to ignore each other during the rest break when they were allowed to leave their isolation rooms. They also were assured that they would *not* meet together or see each other at the close of the experiment.

After the orientation subjects reported to isolation rooms where all subsequent instructions were given at the CRT terminals. The training program was a computerized interactive system which enabled the subjects to move through the material at their own pace. Training involved learning how to operate a computer terminal and complete trade agreements by sending and receiving messages to partners of their choice.

Network Structure and the Bargaining Process

The two networks studied in this experiment are presented in Figure 2a (unbalanced network) and Figure 2b (balanced network). The network in Figure 2c was not included in this experiment. Subjects reported directly to their isolation rooms to start the experimental session. During the first five minutes the subjects engaged individually in a simple symbol-manipulation task in order to acquire through effort the resources to be used in exchange. The resource points earned during this phase were valuable only through exchange.

Transaction tables. Each subject was given a printed *benefit table* which indicated 100 possible trade agreements. The table listed the amount of resources a subject must offer (from 1 to 100) in order to obtain a specified number of benefit points. Each benefit point earned in exchange was worth 3¢ at the end of the experiment. While the subject's own profits were displayed in the table for each offer, the other actor's profits were not revealed. While both members of any negotiating pair had identical printed benefit tables, these were combined in the computer into *transaction tables* having a fixed amount of overlapping joint benefit

¹⁸ Thus, while we have 112 subjects, our N is only 28 for purposes of statistical analysis. The phenomenon under study (see Definitions 3 and 4 above) requires that the entire network be taken as the unit of analysis.

(i.e., a 24-point overlap in A-B and C-C relations and eight points in B-B relations).

The bargaining process. Using the printed benefit table, any subject at any time could send an offer to any one of three alternative trading partners by indicating the number of resource units being offered. The computer would consult the transaction table and inform the other party of the number of units the sender would accept. Thus, with each offer both parties knew their own but not the other's benefit. Before any offer could be accepted, however, two persons had to enter formal negotiation, by one party sending and the other accepting an invitation to negotiate. Thus, the three-minute transaction period typically passes through three phases: (a) all four parties send and receive exploratory *informal* offers; (b) after some time the four people form into two pairs negotiating through *formal* offers; and (c) once any formal offer is accepted a *transaction* is recorded and the S's wait for the next transaction period to begin. However, phase (a) might be short or skipped altogether if the S's move directly into formal negotiation (see *commitment* below); and phase (c) might be missed—with a loss of possible earnings—if S's fail to locate a partner and come to an agreement before the three minutes are up.¹⁹

The power manipulation. The reader should refer to Figures 2a and 2b for a schematic representation of the different network structures created in the experiment by varying the benefit overlaps. The different overlaps (indicated by solid and broken lines in Figure 2) made the potential reward three times greater in A-B and C-C relations (24 benefit points) than in the B-B relations (eight benefit points). Notice that the persons in positions A and

C have identical potential relations with their respective partners. They differ only in the character of the larger structure in which the relations are embedded. Thus, the experiment was designed to compare the outcomes in the A-B and C-C exchange relations. (We have little interest in B-B exchange relations as such in this particular study.)

The Equity Manipulation

The process of completing transactions resulted in the accumulation of benefit for the actors in each network. The subject's own benefit was always visible on the CRT screen after each transaction was completed. At the end of the first 20 transaction periods all persons' cumulative benefits obtained to that point were displayed on all screens for comparison during the ten minute break. Therefore, when negotiations resumed during the second hour of the experiment the subjects would engage in bargaining with knowledge about the relative benefits obtained by each actor in their network from past bargaining. Given knowledge of the distribution of profits across the network normative factors such as concern for equity could then operate to modify former levels of power use.

FINDINGS

Measures were developed and statistical tests were selected in recognition of the following dependencies built into our data. First, when two subjects bargain, the behavior of each is dependent upon the other in a process which results in a single event called a transaction. If subject I gets Y points in a given exchange, then subject J must get $X = (24 - Y)$. Second, our conception of network connection (Def. 4, above) requires that we treat exchange outcomes in different relations as nonindependent events. Third, the concept of comparison level, CL_{ij} , implies that negotiated outcomes are nonindependent through the time series of 40 transaction periods. Thus, the entire four-person network interacting through 40 three-minute transaction periods must form the statistical unit for most of our analyses.

¹⁹ While there have been many studies of the bargaining process (see Rubin and Brown, 1975), the one closest to our own format is Kelley et al. (1967). In that study subjects did not know what benefit the other person made. That study, however, examined only two person bargaining. A study by Stolte and Emerson (1976) used four-person networks similar to those employed here. However, subjects communicated by telephone and were therefore free to (and did in fact) tell one another their respective benefits.

Power and Position in Network Structures

In the two networks under study, some subjects were randomly assigned to the powerful position A in Figure 2a while others were randomly assigned to the power-balanced position C in Figure 2b. These two sets of subjects are assumed to be comparable in all respects other than the structurally determined power attached to their positions. Hypotheses 1 and 2 above predict the distribution of benefits which should be produced through exchanges within these two power structures.

Let CL_{ij} be the average benefit obtained by person I through bargaining with person J over a given time period (say five transaction periods). Since all exchange relations for occupants of both positions A and C have a 24-point benefit overlap, we know that for each exchange $CL_{ij} = 24 - CL_{ji}$. Thus we can measure power use by actor I over actor J as a difference score, $p_{ij} = (CL_{ij} - CL_{ji})$. The basic measure p_{ij} is suitable for aggregating scores for exchange relations between actors who occupy different positions. However, when aggregating p_{ij} scores for exchange relations between actors who occupy the same position, as for the exchange relations among all four occupants of position C, the mean $p_{c_1c_j}$ score is automatically zero due to the character of the difference score. Therefore, our measures of power use will be:

$|P_{c_1c_j}| = |CL_{c_1c_j} - CL_{c_jc_1}|$ for power used within position C; and

$p_{abj} = (CL_{abj} - CL_{bja})$ for power used between positions A and B.

The use of power in balanced networks. We can state hypothesis 1 above in testable form as:

Hyp. 1' Power use in position C will decrease over time toward $|p_{c_1c_j}| = 0$.

Figure 3 presents mean power use, $|p_{c_1c_j}|$, values for each of the eight trial blocks. (Each trial block includes five transaction periods.) Pooling sexes and using a t-test for correlated samples to compare the average of the first two with the average of the last two trial blocks, we

obtain $t = 3.28$ ($df = 13$, $p < .001$). Thus, the hypothesis is supported. However, according to theory, $|P_{c_1c_j}|$ should continue to decline toward zero. Figure 3 suggests that it has stabilized at about 4.0, and probably would not reach zero within the parameters of this experiment. If it has stabilized at 4.0, that value measures a residue of *unused* power in the network.

The use of power in unbalanced networks. Hypothesis 2 can be restated in testable form as:

Hyp. 2' Power use in position A will increase over time toward $p_{abj} = (20 - 4) = 16$.

Figure 4 presents mean power use, p_{abj} , values by trial block. The acquired benefits of both males and females in the powerful position rise rapidly from period 1 to period 2 and continue to climb in periods 3 and 4. After period 4 the relative benefits accumulated were revealed to all subjects thus allowing normative concerns for equity to enter into the subsequent bargaining process. There is no way of knowing how far toward the theoretical maximum of 16 this increase would have gone had we not intervened with the equity treatment. The means and variances of the power use scores across trial blocks are presented in Table 1. Since the experiment consists of 40 transaction periods with the equity manipulation introduced after period 20, transaction periods were divided into eight trial blocks of five periods each (four before and four after the equity manipulation). Table 1 provides a comparison between power use, p_{abj} , in position A and power use, $p_{c_1c_j}$, in position C. Since position A has a single occupant while position C has four occupants interacting with each other, the comparison could be made either by averaging $|p_{c_1c_j}|$ across the four occupants, or by selecting one occupant of position C to represent the network. We chose the latter because it allowed us to use the same algebraic difference score, p_{ij} , to make the comparison.

Table 2 provides the summary statistics for the analysis of variance conducted on the data presented in Table 1. Hypothesis

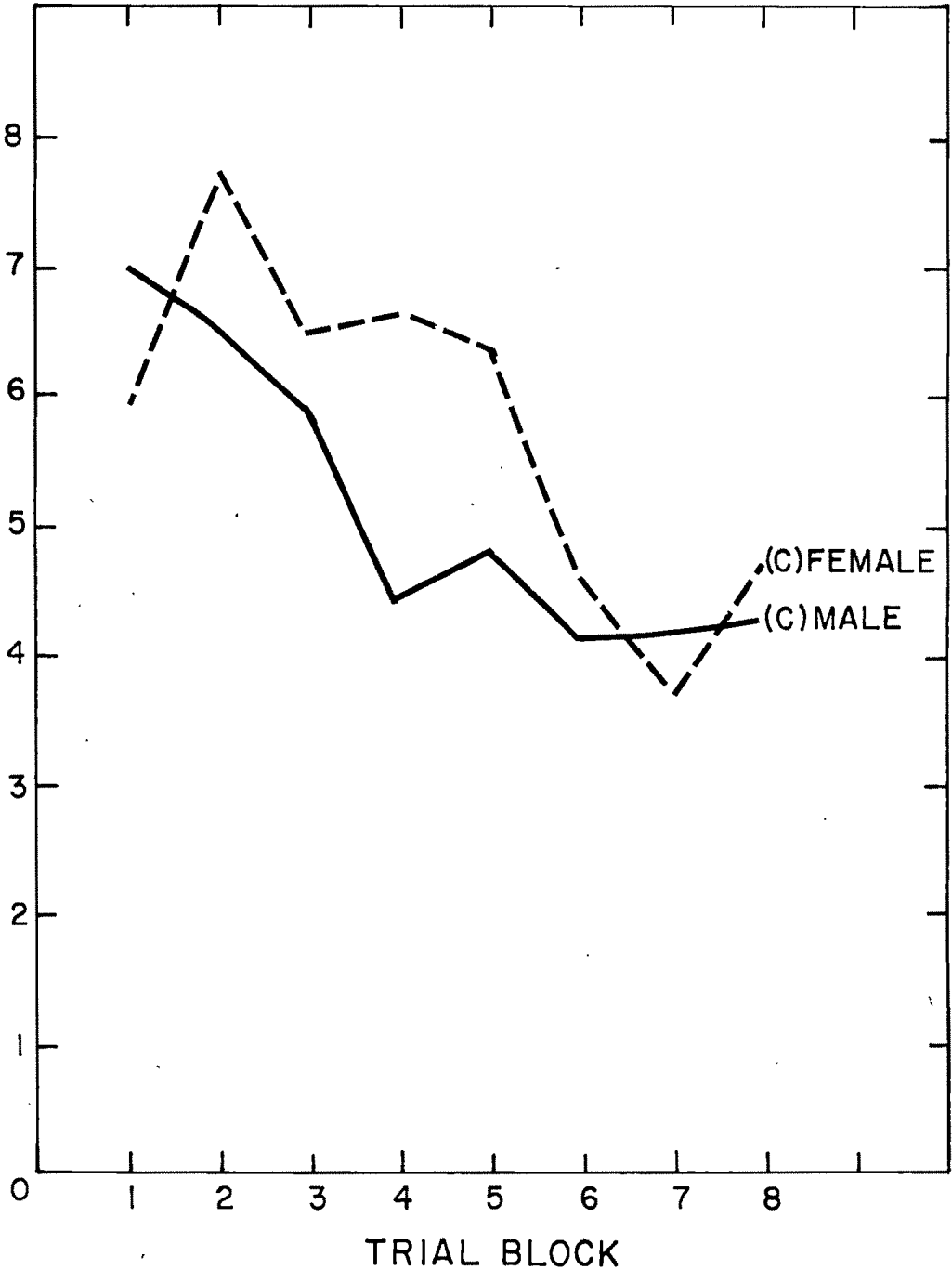


Figure 3. Average Power Use (p_{cc}), in Position C by Sex

2 is strongly supported by the significant main effect of power ($F=10.1$; $df=1, 24$; $p < .001$).²⁰ Power use in position A far

²⁰ This main effect on power is a conservative test of Hypothesis 2. A more direct approach would employ $p_{ab} = 0$ as the null hypothesis for testing Hy-

pothesis 2. However, we prefer to compare power use in position A with power use in position C for conceptual reasons.

exceeds that in position C. (Analysis not reported shows the same results regardless of which occupant of position C is used to represent the balanced network.)

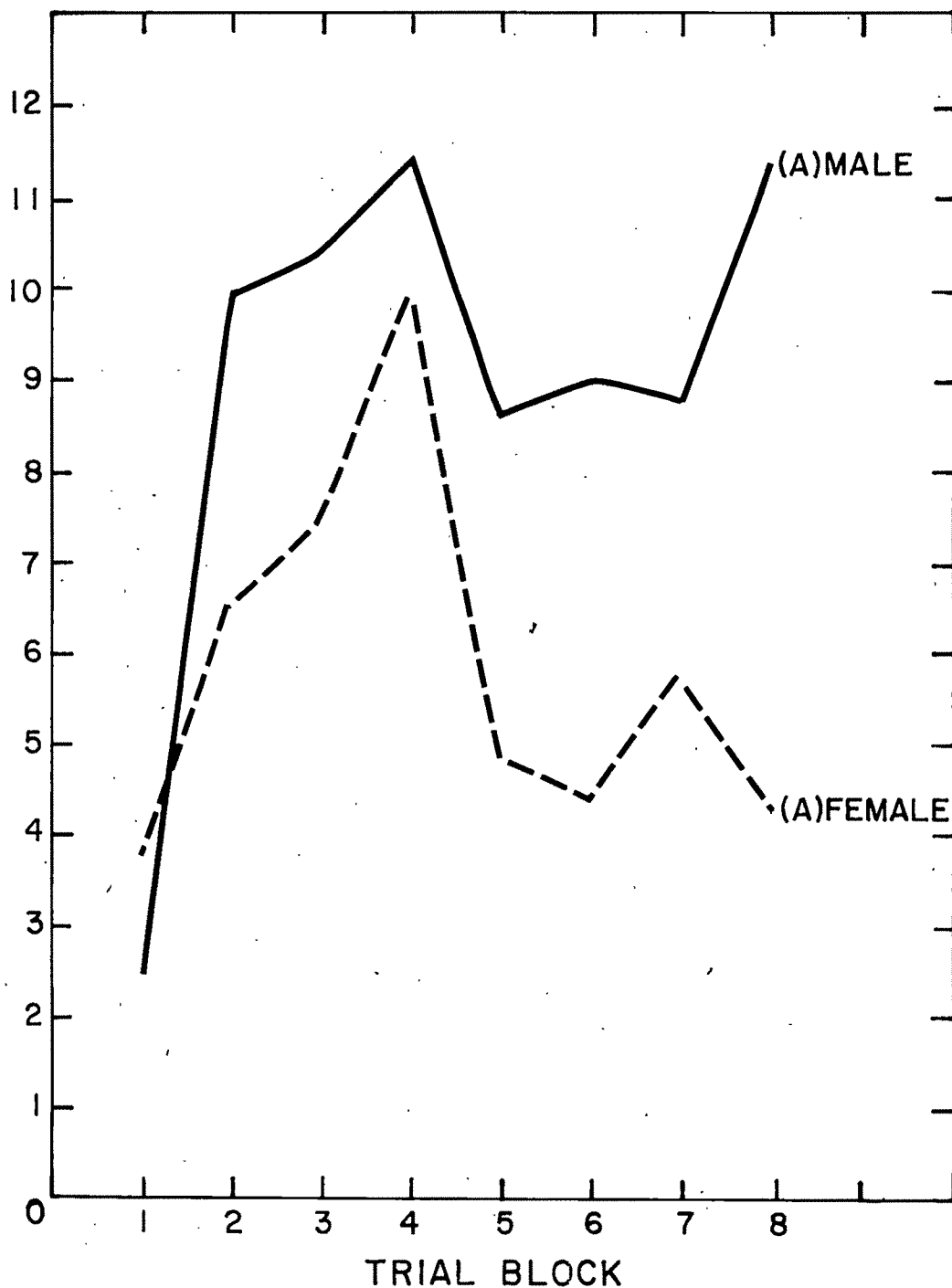


Figure 4. Average Power Use (p_{ab}), in Position A by Sex

Equity Concerns as a Constraint on Power Use

Hypothesis 3 predicted a decrease in p_{ab1} after trial block 4, when each person

was made aware of everyone's cumulated earnings during the first 20 transaction periods. Figure 3 shows a pronounced effect of the kind predicted for both sexes. Having first determined that the sexes were

Table 1. Means and Variance of Power Use, P_{ij} , by Trial Block and Sex for Positions A and C

Sex	Position	Trial Block							
		(Before)				(After)			
		1	2	3	4	5	6	7	8
M	(A)	\bar{X}	2.5	9.9	10.4	11.4	8.6	9.0	8.8
		s^2	60.2	61.8	62.6	41.9	60.3	29.7	35.6
		(n)	(7)	(7)	(7)	(7)	(6)	(7)	(7)
F	(A)	\bar{X}	3.9	6.6	7.7	9.9	4.8	4.4	5.7
		s^2	144.9	97.6	68.7	57.4	126.0	89.7	55.9
		(n)	(6)	(6)	(7)	(7)	(7)	(7)	(7)
M	(C)	\bar{X}	-0.1	-2.2	-2.7	-0.2	.5	1.2	2.9
		s^2	32.1	24.3	25.9	29.7	28.6	26.6	20.5
		(n)	(6)	(7)	(7)	(7)	(7)	(7)	(7)
F	(C)	\bar{X}	.1	5.6	-0.2	2.3	6.1	2.1	-0.4
		s^2	59.4	112.7	85.2	44.7	39.7	48.1	12.0
		(n)	(5)	(7)	(7)	(7)	(7)	(7)	(7)

Note: Only one member of each network is used: the powerful member (A) in unbalanced nets, and an arbitrarily chosen member (C) in balanced nets.

not significantly different in this effect, we pooled sexes and compared trial blocks 4 and 5, the periods immediately before and after the equity treatment. The difference is statistically significant. (Using a t-test for correlated samples, we obtain $t = 1.84$, $df = 13$, $p < .05$.)²¹

Commitment as a Constraint on Power Use

We conceive commitment between exchange partners to be an interpersonal attachment leading persons to exchange repeatedly with the same partners. Some readers might assume that this phenomenon can be measured in two ways: (1) at a behavioral level, in terms of partner choices in a field of alternative partners; and (2) at an attitudinal level through sentiments of loyalty, attraction, liking, etc. *We are not willing to make this assumption* at this early stage in our research. There are theoretical reasons to believe the measures listed above as (1) and (2) involve fundamentally different phenomena related to one another in ways too complex to be dealt with adequately here. We shall define and measure commitment here entirely at the behavioral

level. Commitment so defined is a social-structural phenomenon properly measured as an attribute of total network unit. We reserve for subsequent research the study of attitudinal correlates of position in committed structures. Partner selection in such structures is not a simple matter of choice governed by individual preference. Therefore, we should expect to find complex attitudinal correlates.

The commitment measure. In our four-person networks, after t transaction periods have elapsed, person I will have completed $T_{ij} + T_{ik} + T_{il} \leq t$ transactions with persons J, K and L. We developed a measure of commitment, $[C_i]_t$, for person I after t transactions having a maximum value of 1.0 when (a) I completes a transaction at every opportunity, and (b) has done so always with the same partner. The minimum value of $[C_i]_t = 0$ should be

Table 2. Analysis of Variance of Power Use, p_{ij}

Source	DF	Mean Squares	F	P
Power (P)	1	2263.91	10.1	.001
Sex (S)	1	21.44	.1	NS
P x S	1	348.39	1.5	NS
Within (PS)	24	225.09		
Equity (E)	1	14.39	.2	NS
P x E	1	73.28	1.0	NS
S x E	1	124.07	1.7	NS
P x S x E	1	.11	.0	NS
Within (PS)E	24	71.06		
Trials (T)/E	6	48.18	2.2	<.05
T/E x P	6	44.34	2.04	<.10
T/E x S	6	21.43	.99	NS
T/E x P x S	6	35.62	1.64	NS
Within (PS) x T/E	144	21.72		

²¹ One male A was totally excluded from exchange in trial block 5, thus having no CL score. It is clear that this was a reaction by the others in his network to his use of power, revealed by his benefits displayed after trial block 4. Therefore rather than delete the case he was assigned the mean score of the other 13 subjects for trial block 5.

obtained when I's transactions have been equally distributed among potential partners, so that $T_{ij} = T_{ik} = T_{il}$. The measure is:

$$[C_i]_t = (|T_{ij} - T_{ik}| + |T_{ij} - T_{il}| + |T_{il} - T_{ik}|) / 2t.$$

This measure can be computed for each value of t (after each of the 40 transaction periods), with a cumulating data base which gives the score increasing stability as t increases. While each of the four persons in a network can have a different commitment score, the four scores are obviously interdependent. To the extent that any person becomes committed, others, losing "choice" in partner selection, become structurally committed of necessity.

Hypothesis 4 asserts that power use will vary inversely with commitment. In order to examine this relationship, in Table 3 we present correlation coefficients between power use, p_{abij} , and commitment, $[C_i]_t$, for $t = 5, 10, 15, \dots, 40$ for males and females in the powerful position (A). Dependencies in the data force us to use only 14 units in this analysis, seven male and seven female power unbalanced networks. Furthermore, because of the possible sex differences visible in Table 2 we are reluctant to pool sexes. Therefore, with such small n 's the correlations fall short of significance. Even so, we are encouraged by the direction and pattern of these correlations.

Hypothesis 4 can also be tested through a corollary. *Power use* involves the exploration of opportunities through sending out tentative or informal offers. In our data formal offers almost always went to the party with whom a transaction had been consummated. Therefore, power use through exploring alternatives can be measured for each subject as the propor-

tion of offers made which were informal. We obtained the following correlations between *Commitment* and *% Informal Offers* for $N = 112$ subjects:

Trial Block:	1	2	3	4
r:	-.40	-.34	-.35	-.38

Trial Block:	5	6	7	8
r:	-.33	-.06	-.14	-.23.

A correlation of $r = .25$ is required for significance at the .01 level. (With % Offers Informal as the dependent variable, the subject can be used as the statistical unit.) That these correlations should diminish over time seems reasonable because "exploration" of alternative exchange relations, diminishing over time, reduced the variation to be explained.²²

Sex Differences

While sex is not a variable of major theoretical importance in this study we included it in the experiment in order to examine the sex differences in bargaining behavior alluded to in the literature. Regarding power use, we find no statistically significant sex differences. However, in view of the small n involved, this experiment may provide an inadequate test. The direction of the sex difference depicted in Figure 4 suggests that when in a powerful position females exercise their power less dramatically than do males. When equity concerns become salient after trial block 4, both males and females drop to a lower level of power use, but only the males show a tendency over time to resume vigorous use of power. But again, this difference falls short of significance. Similarly, in the balanced power position, C, females appear to be less forceful than males in their power use; since $|p_{cicj}|$ declines later for females than for males (see Figure 3), and only after relative benefits

Table 3. Correlations between Power Use, p_{ij} , and Commitment, $[C_i]_t$, for Males and Females in the Powerful Position A by Trial Block

Sex (n)	Trial Block							
	1	2	3	4	5	6	7	8
M (7)	-.14	-.48	-.31	-.45	-.45	-.56	-.67	-.56
F (7)	.25	-.17	.23	-.19	.35	-.23	-.48	-.43

²² The relationship between power use and commitment is an interesting and complex question, both conceptually and empirically. The data on power use have not been fully analyzed. However, to carry the analysis further we must conceive power use and commitment in reciprocal interaction through time, and conduct a regression analysis incorporating lagged variables into the model. This type of analysis is being conducted currently as a separate project.

have been revealed. These differences, however, are too small to justify any definite conclusions regarding sex differences in power use.

Sex differences in commitment. The mean commitment scores, $[C]_t$, for males and females on powerful (A) and in power-balanced (C) positions are presented in

Figure 5. Each of the four trend lines is a progression of means for seven cases, each case being a four-person network. A downward trend indicates movement away from previous exchange partners. An upward trend measures the tendency to return to the same partner. As commitment patterns form over time there is a

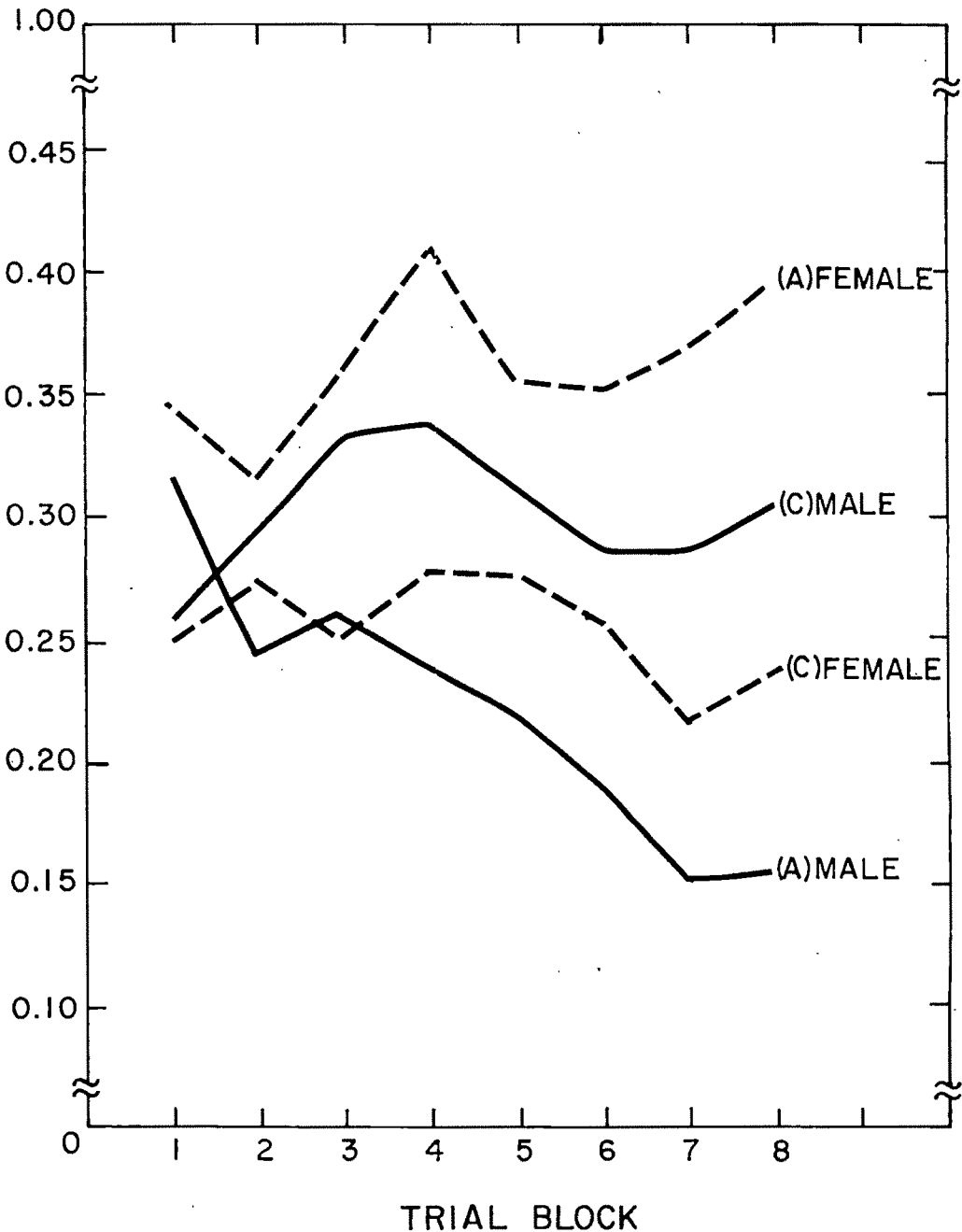


Figure 5. Mean Commitment (C_t), in Position A and C by Sex

pronounced sex difference in the power-unbalanced networks. In the balanced networks, the difference is small and insignificant. The results of a 2×2 (power by sex) analysis of variance conducted on the cumulative $[C_1]_t$ scores at transaction period 40 indicate only a significant interaction effect ($F = 8.03$; $df = 1, 24$; $p < .01$). This finding suggests that in the power-unbalanced networks, females are noticeably more likely to find partners and stay with them than are males.

DISCUSSION AND CONCLUSIONS

Our major purpose in this experiment was to examine structural determinants of power. In comparing the two networks in Figures 2a and 2b, the findings demonstrate the importance of position in a structure as a determinant of power use. Positions A and C in these two networks are proximally identical in all respects; these positions provide the same number of alternatives with the same exchange value, and their occupants (in this experiment) had no information about distal features of the network. However, locations A and C do differ in structural arrangements remote from these locations. We have shown that a developing theory of networks can offer predictions from structure to behavior at different locations in that structure, and our findings support those predictions quite well. These findings also provide evidence that our laboratory setting can be used to study more complex and theoretically interesting structures such as the one portrayed in Figure 2c.

Similarly, our findings suggest that normative concerns operate as constraints upon the use of power in exchange networks. Future experiments will be designed to test predictions derived from competing theories of equity or justice concerning the role of equity concerns in exchange processes in n -person social structures. The findings clearly demonstrate that equity processes and power processes are interrelated in interesting ways. Further research should be conducted concerning the specific character of the link between power and justice processes.

Another major contribution of this

study to social exchange theory is the development of the concept *commitment*, its measurement and the findings relating to it. We suspect commitment will prove to be an important concept separating social exchange theory from strictly economic exchange theory. To the extent that specific social actors are drawn into repetitive exchange with one another (whether through reciprocal reinforcement in casual interpersonal attraction, through institutional arrangements such as marriage or long-term employment contracts, or through collectively or normatively enforced systems of obligation such as kinship systems), the market structure brought to the situation by economic analysis is imperfect and traditional microeconomic theory loses its precision. Needed in place of a theory tied to perfectly competitive (i.e., uncommitted) markets is a theory of network structures tied together by repetitive exchange with specified partners.

Finally, while the sex differences reported here were incidental to our main research objectives, the findings warrant comment. Most important, the major finding, showing a strong sex \times power interaction on commitment, should be interpreted with caution. It would be tempting to invoke the extensive literature which attributes to females a higher level of interpersonal orientation as opposed to the task orientation typically attributed to males. For example, Rubin and Brown (1975:172), in an attempt to reconcile many contradictory studies concerning sex differences in bargaining research, conclude: "our argument is simple, females tend to bargain like high IO's . . . [i.e., high on interpersonal orientation]." But commitment data in the power-balanced networks do not support this argument; females do not exhibit more commitment to their partners than males in position C (see Figure 5). More importantly, the commitment measured here is an emergent structural phenomenon at the network level with potentially complex links back to the attitudes, values or socialized traits of the persons involved in the exchange processes. Before this finding can be fully understood, additional studies are needed which investigate the



determinants of commitment formation in exchange networks. We offer the following hypothesis for future research on this topic:

Commitment formation varies inversely with risk-taking behavior in exchange situations involving uncertain outcomes.

If this hypothesis is confirmed, and if males engage in more risk-taking than do females, then all of the data we have obtained on sex differences can be fully interpreted, including the small differences in power use evident in this study.

REFERENCES

- Abramson, H. A., W. Cutler and M. Mendelson
1958 "Social power and commitment: a theoretical statement." *American Sociological Review* 23:15-22.
- Adams, J. S.
1965 "Inequity in social exchange." Pp. 267-99 in L. Berkowitz (ed.), *Experimental Social Psychology*. New York: Academic.
- Barth, F.
1966 *Models of Social Organization*. London: Royal Anthropological Institute.
- Berkowitz, L. and E. Walster (eds.)
1976 *Equity Theory: Toward a General Theory of Social Interaction*. *Advances in Experimental Social Psychology*, Vol. 9. New York: Academic.
- Bierstadt, Robert
1950 "The analysis of social power." *American Sociological Review* 15:730-6.
- Blau, P.
1964a *Exchange and Power in Social Life*. New York: Wiley.
1964b "Justice in social exchange." *Sociological Inquiry* 34:193-206.
- Burt, Ronald
1976 "Positions in networks." *Social Forces* 55:93-122.
- Coddington, A.
1968 *Theories of the Bargaining Process*. Chicago: Aldine.
- Cook, Karen S. and R. M. Emerson
1977 "Design of a computerized laboratory for the experimental analysis of social interaction." Technical Report, Department of Sociology, University of Washington.
- Cook, Karen S. and T. L. Parcel
1977 "Equity theory: directions for future research." *Sociological Inquiry* 47:75-88.
- Dahl, Robert
1957 "The concept of power." *Behavioral Science* 2:201-15.
- Eder, D., and M. T. Hallinan
1978 "Sex differences in childrens friendships." *American Sociological Review* 43:237-50.
- El-Ansary, Adel and Louis W. Stern
1972 "Power measurement in the distribution channel." *Journal of Marketing Research* 9:47-52.
- Emerson, R. M.
1962 "Power-dependence relations." *American Sociological Review* 27:31-40.
- 1967a "Exchange theory, part I: a psychological basis for social exchange." Pp. 38-57 in J. Berger, M. Zelditch, Jr., and B. Anderson (eds.), *Sociological Theories in Progress*, Vol. 2. Boston: Houghton-Mifflin.
- 1967b "Exchange theory, part II: exchange relations and networks." Pp. 58-87 in J. Berger, M. Zelditch, Jr., and B. Anderson (eds.), *Sociological Theories in Progress*. Boston: Houghton-Mifflin.
- 1968 "Role theory and diminishing utility in group problem solving." *Pacific Sociological Review* 1:110-5.
- 1972 "Exchange theory: the problem of appropriate data." Seattle: Institute of Sociological Research, University of Washington. Presented at the annual meeting of the American Sociological Association, New York, 1971.
- 1974 "Social exchange theory." Pp. 335-62 in A. Inkeles, J. Coleman, and N. Smelser (eds.), *Annual Review of Sociology*, Vol. 2. Palo Alto: Annual Reviews.
- Firth, R.
1967 *Themes in Economic Anthropology*. London: Tavistock.
- Freeman, Linton C.
1977 "A set of measures of centrality based on betweenness." *Sociometry* 40:35-41.
- Homans, G. C.
1961 *Social Behavior: Its Elementary Forms*. New York: Harcourt, Brace.
1974 *Social Behavior*. Rev. ed. New York: Harcourt, Brace.
1976 "Commentary." Pp. 231-44 in L. Berkowitz and E. Walster (eds.), *Advances in Experimental Social Psychology*, Vol. 9. New York: Academic.
- Hopkins, Terence
1964 *The Exercise of Influence in Small Groups*. New Jersey: Bedminster.
- Huesmann, L. R. and G. Levinger
1976 "Incremental exchange theory: a formal model for progression in dyadic social interaction." Pp. 191-229 in L. Berkowitz and E. Walster (eds.), *Advances in Experimental Social Psychology*, Vol. 9. New York: Academic.
- Kelley, H. H., L. L. Beckman and C. S. Fischer
1967 "Negotiating the division of rewards under incomplete information." *Journal of Experimental Social Psychology* 3:361-98.
- Komorita, S. S.
1968a "Power, the use of power, and cooperative choice in a two-person game." *Journal of Personality and Sociological Psychology* 8: 134-42.
1968b "Bargaining and concession making under bilateral monopoly." *Journal of Personality and Social Psychology* 9:15-20.

- Kuhn, Alfred
1964 "Bargaining power in transactions." *American Journal of Economics and Sociology* 23:49-63.
- Leik, R. and S. Leik
1972 "Interpersonal commitment as a balancing mechanism in social exchange." Unpublished manuscript, Department of Sociology, University of Washington.
- Lerner, M. J.
1975 "The justice motive in social behavior." *Journal of Social Issues* 31:1-19.
- Malinowski, B.
1922 *Argonauts of the Western Pacific*. London: Routledge.
- Marks, Stephen R.
1977 "Multiple roles and role strain: some notes on human energy, time and commitment." *American Sociological Review* 42:921-36.
- Meeker, Barbara F.
1971 "Decisions and exchange." *American Sociological Review* 36:485-95.
- Michener, H., E. Cohen and A. Sørensen
1977 "Social exchange: predicting transactional outcomes in five-event, four-person systems." *American Sociological Review* 42:522-35.
- Nagel, Jack H.
1968 "Some questions about the concept of power." *Behavioral Science* 13:129-37.
- Pen, J.
1959 *The Wage Rate Under Collective Bargaining*. Cambridge, Ma.: Harvard University Press.
- Polanyi, K., C. M. Arensburg and H. W. Pearson (eds.)
1957 *Trade and Market in the Early Empires*. New York: Free Press.
- Rubin, J. Z. and Bert R. Brown
1973 "Becoming committed." Pp. 157-81 in Z. Rubin (ed.), *Liking and Loving*. New York: Holt.
- 1975 *The Social Psychology of Bargaining and Negotiation*. New York: Free Press.
- Sahlins, M.
1965 "Exchange value and the diplomacy of primitive trade." Pp. 95-129 in *Proceedings of the American Ethnological Society*. Seattle: University of Washington Press.
- Stolte, J. and R. M. Emerson
1976 "Structural inequality: position and power in network structures." Pp. 131-49 in R. Hamblin and J. Kunkel (eds.), *Behavior Theory in Sociology*. New York: Transaction.
- Thibaut, J. and H. H. Kelley
1959 *The Social Psychology of Groups*. New York: Wiley.
- Walster, E., E. Berscheid and G. W. Walster
1973 "New directions in equity research." *Journal of Personality and Social Psychology* 25:151-76.
- Walster, E. and G. W. Walster
1975 "Equity and social justice." *Journal of Social Issues* 31:21-43.
- Wrong, Dennis
1968 "Some problems in defining social power." *American Journal of Sociology* 73:673-81.

CROSS-NATIONAL DIFFUSION: THE EFFECTS OF CANADIAN TV ON RURAL MINNESOTA VIEWERS*

DAVID E. PAYNE

University of North Dakota

American Sociological Review 1978, Vol. 43 (October):740-756

This paper reports an empirical test of the effect of cross-national television on viewers' attitudes and cognition. Interviews were given to adult (N = 414) and high school (N = 218) populations in three northern Minnesota locations; one receiving Canadian TV only, one receiving Canadian and U.S. TV, and one receiving U.S. TV only. Controls for other media use, other exposure to Canada, and various socioeconomic dimensions were employed. Substantial cognitive and attitudinal effects were found. Cognitive effects were generally stronger than attitudinal effects, and cognitive and attitudinal effects toward sending country were stronger than toward receiving country. The pattern of effects differed somewhat in the adult and student samples.

This paper examines some of the effects of TV in one of the three commonly used settings. Most commonly, TV effects are examined in an *intranational* setting. Extensive literatures using this setting exist for many types of effects including both knowledge of political and cultural issues and feelings about them. Despite the abundance of research in this setting, the number of firm conclusions remains limited. One reason for this is the ever presence of TV. Most people in the U.S., and Western nations generally, have had access to TV for years. This makes it difficult to find non-TV viewing control populations which are not otherwise atypical. Experimental studies which expose TV-wise subjects to a particular type of TV programming for a few hours or even a few weeks may find little effect against a background of years of exposure.

To overcome this serious problem and at the same time to obtain information which could have significant applied impact, many researchers have turned to

studying the effects of TV programming (usually imported) in less developed countries. Schramm (1964), Rogers (1962), Contreras et al. (1976), Arnove (1975), Horowitz (1975), Schiller (1974), Almaney (1972), and many others have reported substantial effects. Indeed, anecdotal and case study reports of the effects of TV on developing countries go back at least 20 years. Areas of focus have ranged from basic cultural values (Tsai, 1970) to scores on psychological intelligence tests (Madigan and Peterson, 1977).

One serious limitation of this second setting for research is the confounding of technological with ideological or cultural changes. Effects on viewers' knowledge of, and attitude about, cultural and political items may occur independently or to some degree through adoption of technological items. The questions of whether the same effect would occur independent of technological differences in the cultures is unanswered.

A third setting for examining the effects of TV is becoming popular. Cross-national effects of TV between *developmentally similar* cultures are examined. Thus, ideological, political, and cultural effects can in some measure be separated from technological differences. Much of this research has come as a result of concern for the effects of U.S. TV on Canadian viewers (Beattie, 1967; Davey, 1972; Elkin, 1972; Skipper, 1975; McCombs, 1976b). Research also is appearing, however, on the effects of U.S. TV on other nations (Payne and Peake, 1977; Payne, 1978;

*Address all communications to: David E. Payne; Department of Sociology; University of North Dakota; Grand Forks, ND 58201.

Partial funding for this project came from a University of North Dakota Faculty Research Grant and a grant from the Newhouse School of Public Communications, Syracuse University. I appreciate the comments of Gordon Whiting, Charles Mueller, William Fox, Edward Lawler, Maxwell McCombs, Joe Dukim and anonymous *ASR* reviewers who read an earlier draft. I also appreciate several graduate students at the University of North Dakota who spent a great deal of time, without much financial return, gathering data.

Varis, 1974; Read, 1976) and the effects of other countries' TV on U.S. viewers (Sparkes, 1977).

One of the problems that has characterized research in both the less developed nation and the developmentally similar nation settings is the lack of systematic, consistent findings across studies. One obvious reason for this is that cross-cultural TV may have effects of differing strengths on different attitudinal and cognitive dimensions of the receivers. Tsai (1970), studying the effects of U.S. TV in Formosa found, for example, no change in viewers' fundamental values but some change in specific attitudes toward U.S. culture. Madigan and Peterson (1977: 185) found that Eskimo children having their first exposure to TV, when compared with a control group, showed an increase in "certain kinds of auditory information processing . . . and verbal intelligence . . . [but] no improvement . . . when the questions required verbal responses. Tests of educational achievement [also] did not show a difference. . . ."

Findings are often contradictory, however, even when the same or similar dimensions are studied (Robinson, 1972). Caron (1977) found, for example, increasing favorableness of attitudes toward the country which the TV programming was about, but Tsai (1970), using a different sample, found some evidence for decreasing favorableness of similar attitudes. Lin and Burt (1975) found small behavioral but large information transfer effects, but Payne and Peake (1977) found limited information transfer effects, and McCombs (1976b) found only very slight agenda-setting effects.

What this illustrates is the need for a more systematic, *simultaneous* analysis of the effects of cross-cultural TV on a *number* of clearly different dimensions. This multivariable framework could then be applied in several different cross-cultural settings to assess comparability of findings in their different settings, facilitating formation of a general theory. Contreras et al. (1976) support this position and suggest categories of language, psychological, and political effects. While these categories are interesting, they do not appear to be systematic but illus-

trative. Language acquisition is one aspect of the broader category of information acquisition or, perhaps still broader, cognitive effects. Attitudinal effects reflecting positive or negative dispositions to the existing systems or parts of existing systems (Schiller, 1973) are suggested here as a second basic category. The third dimension which suggests itself is the behavioral dimension.

In addition to the distinction between cognitive, attitudinal, and behavioral categories, a second distinction is theoretically necessary. Cross-national TV can affect performance on any of these dimensions as they relate to the country of residence or the country from which the broadcast is made. U.S. viewers of Canadian TV can have either more or less favorable attitudes toward Canada and more or less favorable attitudes toward the U.S. Thus, Caron's (1977) finding that Inuit children exposed to foreign TV are more favorable to foreign nationalities, may not be particularly disturbing as long as the favorableness toward their own country remains undiminished.

In this paper I examine the effects of cross-national TV viewing on several cognitive and attitudinal dimensions, both as these relate to the country from which the signals are sent and the country in which the viewers reside. The guiding orientation is not a particular theory but an eclectic attempt to provide a broadly conceptualized data base to serve as a first step in constructing a more general theory of mass-media effects. This paper focuses on TV in developmentally similar cultures. Only limited inferences are suggested for developing nations and for other mass media.

Watching foreign TV has been found to increase some types of knowledge about that foreign country (Beattie, 1967) and intuitively that makes sense. If this viewing replaced viewing of the local TV, it also would be expected that knowledge of the local country would decrease. If, on the other hand, it was an alternative to *no* TV viewing, or if it contained much local information, it might actually increase knowledge of one's own country, unless it caused a reduction in the use of other local media (Lin and Burt, 1975). In this case,

Canadian TV contains a great deal of news about the U.S. and up to 36% U.S. programming (Davey, 1970), but the comparison population has open access to U.S. TV and therefore would have available more U.S.-related information. I proposed as a first working hypothesis, therefore, that watching Canadian TV would increase American viewers' information about Canada but decrease their information about their own country.

Although Sparkes (1977) found no relationship between viewing foreign TV and attitude toward the sending country, both Payne (1977) and Caron (1977) found an increasing favorableness toward the (foreign) sending country. In each of these cases the content of the programming was controlled and predominantly favorable to the sending country. This should, as hindsight suggests, have been considered in formulating a second hypothesis. Nordenstreng and Varis (1973) maintained that capitalist TV should reduce favorableness of attitude toward more socialist *receiving* countries, and Beattie (1967) and Schiller (1974) indicated that foreign TV would impede the growth of national (receiving country) identity. As a second working hypothesis I proposed, therefore, that respondents' attitude toward the *sending* country would become more favorable and their attitude toward (identification with) their own country would become less favorable as a result of watching foreign TV.

In agreement with the work of Lin and Burt (1975) and Tsai (1970), I proposed as a third working hypothesis that information acquisition would be more affected than attitude acquisition by viewing foreign TV. The three working hypotheses, then, deal with cognitive effects, attitudinal effects, and the comparative strength of the two.

METHOD

One of the most serious problems of most research dealing with the cross-national effects of TV in developmentally similar cultures is self-selection. Respondent viewers can choose either national or foreign TV. Thus, the foreign TV might produce peculiar attitudes in its viewers or

be watched because they already have those attitudes.

The problem of self-selection was overcome in this research by isolating a small area of north central Minnesota which, because of its geographic location, was able to receive Canadian but not U.S. TV. I demographically matched (age, income, occupation, education, religion, ethnic origin) this area with one area in northern Minnesota which received U.S. and Canadian TV and one which received only U.S. TV.¹ Thus, people did not choose which nationality of TV they would watch; geographic location made the decision.

I constructed and pretested an interview schedule on a college student and a northern Minnesota adult sample. This was then given to one adult in each household of our three selected areas (complete enumeration of households). Useable completion rates were: Canadian TV only area, 92% (N=199); U.S. and Canadian TV area, 90% (N=100); U.S. TV only area, 90% (N=115). There were 414 total useable adult interviews for an average return rate of 90.5%.

In addition, interviews were given to all ninth, tenth, and eleventh grade students in attendance on a given day at four area high schools.² Eighty-eight percent of the students in the school on these days completed useable schedules (N=280). One hundred thirty-two students lived in areas receiving only Canadian TV, 101 received both U.S. and Canadian signals, and 47 received only U.S. signals. Students were given the instrument in small groups during the school day.

Although in each area the same nationality of TV had been available for at least ten years, the possibility of immigration caused some concern as to the length (or strength) of manipulation. Less than a dozen respondents reported moving into the Canadian TV only area in the last year, however. I therefore concluded that this was not a significant disturbing influence.

¹ Details of the demographic matching of the samples and the exact location of the research sites are found in Schulke (1977) and are available on request.

² In one high school over half of the sophomores were at a band contest.

Table 1. Correlations between Independent Variable Measures

Question number	Adult Population			High School Population		
	1	2	3	1	2	3
1. Geography	—			—		
2. Percent of TV time viewing Canadian TV	.83	—		.82	—	
3. Percent of TV news time viewing Canadian TV News	.79	.83	—	.76	.84	—

Data were gathered by ten graduate and undergraduate interviewers. Interviewers were given written and verbal instructions and conducted mock interviews before going into the field. In addition, they met daily with the project director during data collection to discuss possible problems. The interview schedule took 25 to 45 minutes to complete and was the same for the adult and high school populations with the exception of necessary modifications in the demographic questions appropriate to the age group. Data were gathered during a two-week period in April, 1977.

MEASURES

The interview schedule contained three measures of nationality of TV viewed. The first measure (geography) was simply physical location. Respondents were coded as to whether they lived in an area where only U.S. (1), both U.S. and Canadian (2), or only Canadian (3) TV signals were available.³ This ordinal indicator measures the amount of Canadian TV available.

There was, of course, some variation in viewing pattern in the area where both U.S. and Canadian TV signals were available. The first indicator of TV viewing (geographic location) grouped all these together; it did not measure the variation among the respondents where both signals were available. While this did reduce the variation, it eliminated problems of self-selection of signal and inaccurate perception or reporting of viewing patterns.

³ This classification was accurate in all but half a dozen cases. In these cases respondents living in the Canadian TV only area reported that with special antennas they were able to receive U.S. TV occasionally. Only one was able to get both U.S. sound and picture while the interviewer was present, however.

In order to better assess the dual reception area, we included a question asking respondents what percent of their viewing time was spent watching Canadian TV and a question asking what percent of their TV news viewing time was spent viewing Canadian TV news. These two interval level indicators also were used as measures of cross-national viewing (independent variable). Having both a measure of general Canadian TV viewing and a measure of Canadian TV news viewing allowed me to assess the relative effect of news and nonnews programming on foreign TV.⁴

Responses for interviewees in the Canadian TV only and U.S. TV only areas were the same on the second and third measures as for the first measure, since only one signal was available. Both these measures, in addition, however, allowed for variation in the dual reception area. While these items measured variation in the dual reception area, they also allowed interviewee judgment bias and self-selection bias in responses from that area. Nevertheless, they provide a useful cross check on the first item. The high interitem correlation (Table 1) between these measures is expected and gives evidence of all the independent variable measures' reliability.

In specifying dependent variables, I first distinguished cognitive effects from attitudinal effects, then further subdivided indicators into effects relating to the country in which the viewers resided, those relating to the country sending the TV signals, and those relating to both countries.

Choosing representative measures from the multitude of possible cognitive and at-

⁴ This distinction proved to have no effect in the actual analysis.

titudinal dimensions posed a basic dilemma. Not all dimensions could be covered or the interview schedule would be untenably long, but several differing dimensions had to be sampled or the research would have limited comparability value. Both general and specific, political, and nonpolitical indicators of attitudes and cognitions toward sending and receiving country were necessary.

Assessing cognitive effects, I first measured knowledge of issues in the sending country, the receiving country, and in both countries. Respondents were asked to name the most important issues facing the U.S., Canada, and those facing the U.S. and Canada together. This was a general level question which could have evoked political or nonpolitical responses. As a simple measure of the amount of information known about the U.S. and Canada, I first coded only the number of issues listed in each case. I followed each of these questions with a question asking which issue was most important, then three questions asking respondents to indicate some of the facts of the issue, some of the people involved in the issue, and some of the solutions to the issue. Responses to these last three questions were each coded as having no knowledge (usually stated by the respondent) or having some knowledge; then they were additively combined to form indicators of knowledge of the U.S. issue and knowledge of the Canadian issue.

As a specific cognitive measure, respondents were asked to provide the American equivalent of seven Canadian terms, some political (premier, Dominion Day, packs or paki) and some nonpolitical (chesterfield, grey cup, serviette, and back bacon). Factor analysis (Varimax rotation) of these responses indicated a single dominant factor. From this I formed a scale measuring linguistic familiarity using factor loadings (Nie et al., 1975). Two very small subfactors also appeared, one loading on the political and the other on the cultural items. This distinction was empirically unimportant when compared with basic knowledge or lack of knowledge of the language dimension.

There are, then, as indicators of cognitive effect, general questions measuring

knowledge of political and nonpolitical issues relating to sending country, receiving country, and both countries; and specific knowledge questions measuring language transfer.

It seemed much more difficult to sample the large variety of possible attitude transfers. Many of the attitudinal items used in this study were taken from questionnaires developed at Syracuse University by McCombs and Sparkes and employed on their Kingston, Ontario, and upstate New York studies. Their scales had been adapted from other cross-cultural research and had been successfully employed in the U.S.-Canadian setting. I included McCombs's indices of general attitudes toward Canada and toward the U.S. (each is a five-item scale). This was intended to measure shifts in national identification on the part of the respondents which Schiller (1974) and Beattie (1967) both suggested would be an important dimension. I also included abbreviated versions of their scales measuring attitudes toward two ethnic groups in Canada and in the U.S. These are cultural rather than political in orientation and were intended to measure favorableness toward the people of the two countries.

Two types of agenda-setting measures (open-ended and paired comparison) were also included. The open-ended questions asked respondents to specify what they thought was the most important issue facing the U.S., Canada, and the U.S. and Canada together. The paired comparison questions tied the respondents to more specific issues. It contained all possible pairs of five issues (economic recovery, international peace, pollution, energy, and taxes) and asked respondents to choose from each pair which issue was most important for the U.S. today. I interpreted these questions as more attitudinal than cognitive because they involved opinion rather than simply recall.

Still more specifically, separate questions were included on several cultural and political issues. Knowing that Canadian TV has more nudity and less violence than U.S. TV (Strode, 1978; Singer, 1971), I included an item asking respondents whether they felt nudity was more acceptable than violence on TV. I thought that

Table 2. Dependent Variable Measures, Number of Indicators, Question Style, and Reliabilities

Question Cognitive Measures		Question Style	Alpha	
			Adult	High School
Relating to sending country				
Number of issues facing Canada	1	open ended	—	—
Knowledge of Canadian issues	3	open ended	.77	.74
Knowledge of Canadian language	7	open ended	.69	.68
Nudity on Canadian TV	1	likert	—	—
Relating to receiving country				
Number of issues facing U.S.	1	open ended	—	—
Knowledge of U.S. issues	3	open ended	—	—
Relating to both countries				
Number of issues between U.S. and Canada	1	open ended	—	—
Affective Measures				
Relating to sending country				
Attitude toward Canada	5	likert	.59	.64
Attitude toward English Canadians	4	likert	.74	.82
Attitude toward French Canadians	4	likert	.75	.82
Most important issue facing Canada	1	open ended	—	—
Open-ended Canada agenda item	1	open ended	—	—
Relating to receiving country				
Attitude toward U.S.	5	likert	.65	.67
Attitude toward U.S. white	4	likert	.70	.79
Attitude toward U.S. black	4	likert	.77	.86
Most important issue facing the U.S.	1	open ended	—	—
Ranking U.S. issues	10	paired comparison	—	—
Open-ended U.S. agenda item	1	open ended	—	—
Relating to both countries				
Nudity is less harmful than violence	1	likert	—	—
Pro socialized medicine	1	likert	—	—
Pro govt. housing support	1	likert	—	—
Most important issue between Canada and U.S.	1	open ended	—	—
Open-ended U.S. and Canada agenda item	1	open ended	—	—

viewers of Canadian TV would, as a result of their viewing exposure, be more favorable (accustomed) to nudity and less to violence on TV than viewers of U.S. TV.

Following Peers's (1972) conclusion that Canadian TV material is less capitalistic than U.S. TV material and more concerned with national interest or opportunity equalization, I asked one question measuring favorableness toward a government-controlled medical system and one measuring favorableness toward government support of low cost housing.

Thus, the measures of attitudes included national identification scales, ethnic favorableness scales, agenda-setting items, and items relating to specific cultural and political issues. These measures were at both the general and specific levels.

All dependent variable measures are arranged by category in Table 2. Table 2 also shows the number of indicators for each construct where multiple indicators are used and the question style and the alpha reliability coefficients for multiple item constructs.⁵

One of the major flaws of much previous research has been the failure to control for contaminating influences. I dealt with this problem by employing 24 control variables from three major categories, each of which has been shown to influence media effects: other media use (frequency

⁵ It is not possible to include the exact wording in all these items because of space constraints. However, an idea of their style and content can be obtained from Table 2 and a copy of the questionnaire is available on request. No measures of the behavioral dimension were made.

of reading newspapers, time spent reading newspapers, time spent reading Canadian newspapers, time spent listening to the radio, frequency of listening to Canadian radio, frequency of reading magazines, time spent reading magazines, time spent reading Canadian magazines or reading magazines about Canada, and interest in local, national and world news); other Canadian exposure (immediate family in Canada, other relatives in Canada, friends in Canada, and frequency of visiting Canada); and status (age, sex, marital status, ethnic background, education, occupation of household head, family income, religion and religious activity).

Two of the three independent variable measures were interval, the other ordinal. The multiple-indicator dependent variable constructs were typically considered interval (Kim, 1975; Labovitz, 1967;

1970). Following Labovitz's and Kim's suggestion, I employed interval level statistics throughout the analysis. This allowed a much more parsimonious reporting of the complex multivariate analyses, facilitated comparability, and allowed me to employ all the control variables.

I first calculated zero-order correlations. Then, using a step-by-step regression for each relationship contained in Tables 3 and 4, I introduced the 24 control variables. While it is not possible here to show all the combinations of controls employed for each relationship (I examined 24 combinations for each relationship), the 24th order partials for these relationships are found, as well as the zero-order correlations, in Tables 3 and 4. Where lesser order partials result in substantial changes, they are described in the text.

Table 3. Pearsonian Correlation Coefficients between Viewing Canadian TV and Measures of Cognitive Effect

Cognitive Indicator	Geographic Location		% of Time Watching Canadian TV		% of TV News Time Watching Canadian TV News	
	Zero Order	With Controls	Zero Order	With Controls	Zero Order	With Controls
Adult Sample						
Relating to sending country						
Number of Canadian issues named	+.232	+.203	+.249	+.215	+.232	+.216
Ability to provide information about most important Canadian issues	+.330	+.327	+.345	+.326	+.334	+.319
Knowledge of Canadian language	+.510	+.446	+.457	+.401	+.404	+.400
Relating to receiving country						
Number of U.S. issues named	-.060	-.014	-.057	-.056	-.128	-.092
Ability to provide information about most important U.S. issues	-.088	-.087	-.025	-.021	-.079	-.107
Relating to both countries						
Number of issues named between Canada and U.S.	+.111	+.123	+.175	+.183	+.151	+.167
High School Sample						
Relating to sending country						
Number of Canadian issues named	+.307	+.252	+.222	+.150	+.234	+.248
Ability to provide information about most important Canadian issues	+.354	+.322	+.257	+.178	+.258	+.249
Knowledge of Canadian language	+.337	+.182	+.236	+.137	+.258	+.157
Relating to receiving country						
Number of U.S. issues named	+.118	+.064	+.038	-.004	+.010	+.004
Ability to provide information about most important U.S. issues	-.046	-.049	-.041	-.052	-.066	-.034
Relating to both countries						
Number of issues named between Canada and U.S.	+.220	+.224	+.118	+.010	+.120	+.148

COGNITIVE EFFECTS

The zero-order correlations between nationality of TV viewed and various measures of cognitive effects are reported in Table 3. Correlation coefficients, but not significance levels, are reported since three populations rather than samples were enumerated.

There appeared to be almost no relationship in either the adult or high school population between cross-national TV viewing and number of U.S. issues respondents could name, or the amount of information they could provide about those issues. Two of 12 correlations were greater than .10 but they were in opposite directions and both dropped to less than .10 when controls were added. These data fail to support the concerns of some mass communications researchers and politicians that cross-cultural TV results in the loss of one's own culture. Local information contained on Canadian TV and other sources filled the gap.

When this is compared with the correlation between Canadian TV viewing and number of *Canadian* issues respondents could name, and information the respondents could provide about those issues, a dramatic contrast is provided. Both high school and adult respondents who watched Canadian TV were able to list more issues and to provide more information about them than were viewers of U.S. TV. None of the controls or any stepwise combination of them reduced any correlation coefficient for the adult or high school population by as much as .1. The small reductions that did occur were, for the adult population, mostly the result of controlling on frequency of visiting Canada, and, for the high school population, the result of controlling on frequency of listening to Canadian radio.

The greater effect of the control variable, listening to Canadian radio, on high school rather than adult respondents probably results from their greater frequency of listening to Canadian radio. Fifty percent of the high school student population reported listening to Canadian radio daily while only 27% of the adult sample did. The greater effect on adults of

visiting Canada appears to be not so much the result of frequency of visits to Canada in a given time period, since they actually reported visiting Canada slightly less frequently during a given year than the high school sample, but the result of the accumulated effect of many years of visiting.

When the scale of language transfer was correlated with viewing of Canadian TV, the strongest cognitive effects appeared. Those who watched Canadian rather than U.S. TV were more likely to be able to provide the American equivalent of Canadian words. It is interesting to note that the effect was stronger in the adult than in the high school population. This is probably the result of greater length of exposure of adults to Canadian TV. None of the control variables substantially changed these correlations in the adult population. In the high school population the addition of the control *frequency of listening to Canadian radio* reduced the correlations by almost half. No other controls markedly changed the relationship. This indicates that both Canadian TV and radio had an effect on language transfer.

I also correlated viewing of Canadian TV with the number of issues between the U.S. and Canada that respondents could name. These correlations, as expected, fall in every case between those relating to identifying the U.S. and identifying the Canadian issue. It thus appears to follow the theory that viewing Canadian TV as compared with U.S. TV provides viewers with more information about Canada and to a lesser degree external issues with which she is involved. The addition of the 24 controls did not alter substantially these correlation coefficients.

Generally, I found watching Canadian TV was moderately associated with knowledge of Canada, and Canadian-U.S. relations, but not associated with knowledge of the U.S. and U.S. issues. The effect of controls on the association between Canadian TV viewing and the cognitive measures was negligible on all but one variable on one of the two subsamples (language transfer for high school students). Even on this variable some relationship remained after employing 24 controls.

ATTITUDINAL EFFECTS

The correlation coefficients relating viewing of Canadian TV to the attitudinal measures are contained in Table 4. While the cognitive effects on adults and high school students were similar, the attitudinal effects are somewhat different. I will, therefore, deal first with the adults, then the high school students.

Adult Population

Correlating viewing of Canadian TV with favorableness of attitude toward the U.S., attitude toward U.S. whites, and

attitude toward U.S. blacks indicated only very minor relationships. These fluctuated unsystematically from direct to inverse, further indicating the lack of a meaningful association.

Introduction of all 24 controls makes a substantial change for only the relationship between attitude toward U.S. blacks and viewing Canadian television. Those who watched Canadian TV were slightly more favorable to U.S. blacks than those who watched U.S. TV. This provides some evidence for the contention that U.S. TV portrays a negative image of blacks (Rivers, 1972). Even in this case, however, the relationship is slight and not

Table 4. Correlation Coefficients between Viewing Canadian TV and Measures of Attitudinal Effect

Attitude Indicator Adult Sample	Geographic Location		% of Time Watching Canadian TV		% of TV News Time Watching Canadian TV News	
	Zero Order	With Controls	Zero Order	With Controls	Zero Order	With Controls
Relating to sending country						
Attitude toward Canada	-.130	-.132	-.120	-.126	-.144	-.157
Attitude toward French Can.	-.069	-.042	-.063	-.039	-.085	-.077
Attitude toward English Can.	+.008	-.004	-.026	-.011	-.040	-.040
Most important issue facing Canada	+.340	+.320	+.292	+.302	+.362	+.364
Relating to receiving country						
Attitude toward U.S.	-.019	-.028	-.017	+.002	+.004	-.005
Attitude toward U.S. white	+.010	+.028	+.089	+.047	+.044	+.035
Attitude toward U.S. black	-.035	-.095	+.078	+.198	+.061	+.146
Most important issue facing U.S.	-.009	+.018	-.044	-.039	-.061	-.058
Relating to both countries						
Nudity is less harmful than violence	+.121	+.069	+.081	+.056	+.099	+.063
Pro socialized medicine	+.034	-.026	+.048	+.017	+.079	+.058
Pro govt. housing support	+.022	+.033	-.000	+.025	+.019	+.047
Most important issue between U.S. and Canada	+.116	+.099	+.148	+.146	+.199	+.211
High School Sample						
Relating to sending country						
Attitude toward Canada	-.386	-.419	-.279	-.304	-.327	-.340
Attitude toward French Can.	-.146	-.156	-.037	-.070	-.070	-.062
Attitude toward English Can.	-.089	-.153	-.080	-.115	-.036	-.062
Most important issue facing Canada	+.347	+.314	+.243	+.197	+.272	+.273
Relating to receiving country						
Attitude toward U.S.	-.102	-.038	-.022	+.013	-.032	+.028
Attitude toward U.S. white	-.155	-.147	-.072	-.048	-.087	-.062
Attitude toward U.S. black	+.102	+.084	+.050	+.033	+.100	+.089
Most important issue facing U.S.	-.036	+.015	-.173	-.166	-.090	-.089
Relating to both countries						
Nudity is less harmful than violence	-.131	-.128	-.104	-.061	-.181	-.164
Pro socialized medicine	-.124	-.130	-.085	-.094	-.116	-.120
Pro govt. housing support	-.027	-.154	+.027	-.129	+.008	-.156
Most important issue between U.S. and Canada	+.361	+.346	+.238	+.237	+.219	+.253

consistent across all three independent variable measures.

The question asking the respondents to indicate the most important issue facing the U.S. was negligibly related to viewing Canadian TV. Thus, viewing of Canadian TV does not appear to markedly effect the attitudes of adult American viewers about their own country or white or black groups in their country or the selection of most important issue for their country.

Attitudes toward French and English Canadians were not substantially related to viewing Canadian TV for adults, though the direction of the relationship is fairly consistent (those who view Canadian TV are *slightly* less favorable to both French- and English-speaking Canadians). This fails to give support to the concerns of some Canadian media researchers that English language CBC broadcasts create a negative image of French-Canadians. (Dunton and Laurendeau, 1965:73). This finding contrasts with the slight relationship between viewing U.S. TV and negative attitude toward U.S. blacks discussed earlier. Viewing of Canadian TV was related, however, to attitude toward Canada: Adults viewing Canadian TV were *less* favorable to Canada than those viewing U.S. TV. This relationship is slightly strengthened by the introduction of the control set.

While selection of most important issues facing the U.S. was not related to viewing U.S. TV, selection of most important issues facing Canada was. Those who watched Canadian TV were more likely to choose economic stability and less likely to choose international peace than viewers of U.S. TV. This probably results from the economic difficulties that Canada was undergoing at the time and the fact that their country is less involved as an international power than the U.S. These correlations were unaffected by the introduction of controls. Viewing of Canadian TV was related also to choice of the most important issue between the U.S. and Canada. The relationship was weaker but was unreduced by controls.⁶

⁶ Eta measures of these correlations are found in Payne (1977). Although they are slightly stronger, indicating some curvilinearity, the basic pattern of relationship is unchanged.

As a second technique of assessing the agenda-setting function of cross-cultural TV viewing, I used a set of questions pairing all possible two-item combinations of five issues. A technique was employed (Guilford, 1954) which places items on a metric, as well as orders them. These data are contained in Figure 1 and are rather striking. While pollution and taxes are least important for all groups, concern with international peace moves from most important for viewers of U.S. TV to only middle importance for Canadian viewers, and declines regularly in its position on the metric.

These findings give some indication that the media do, in some cases, have an agenda-setting effect as McCombs (1976a) and many others have suggested.

When attitudes toward government involvement in housing and medical care were correlated with Canadian TV viewing, the association was negligible and unchanged by the introduction of controls. There was a slight relationship between viewing Canadian TV and preference for sex over violence on TV. This may be the result of the greater prevalence of nudity and the lesser prevalence of violence (Schulke, 1977; Singer, 1971) on Canadian than U.S. TV. Adults were more accepting of what they had become accustomed to seeing on TV. This relationship is, however, very small.

High School Population

The correlations between viewing Canadian TV and attitudes about the U.S. were stronger, but less stable for the high school than the adult population. The inclusion of the controls reduced all these correlations, in the high school population, making them more similar to those found for the adult population. It appears that for high school students there are no substantively important relationships between watching Canadian TV and evaluation of the U.S., U.S. whites, or U.S. blacks.

High school interviewees who viewed Canadian TV, like their adult counterparts, showed only slightly less favorable attitudes toward French- and English-speaking Canadians and these were simi-

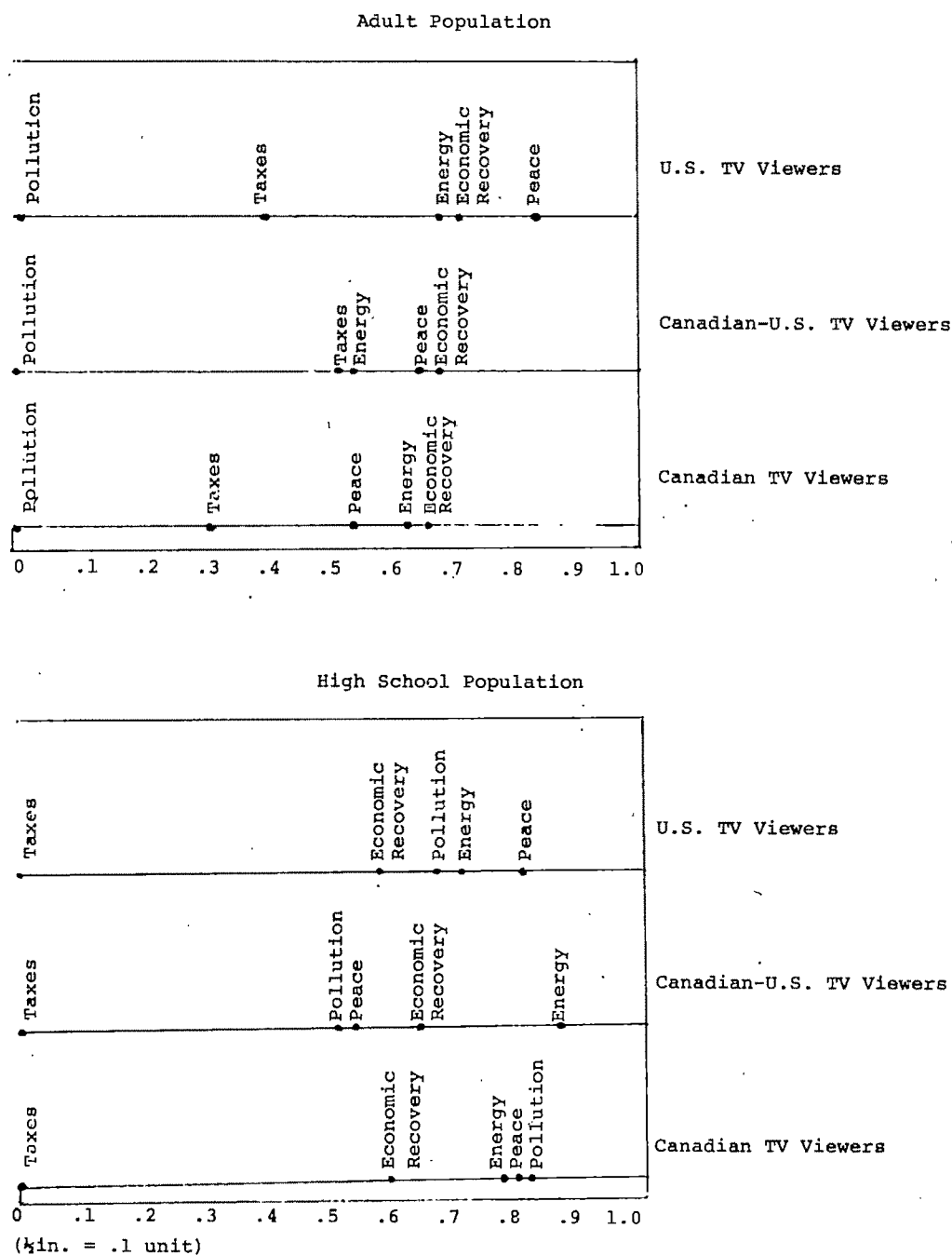


Figure 1. Scales of Mean Z-Scores from Paired-Comparisons

larly unfavorable toward both. The strength of the relationship was increased slightly, however, when controls were added. The most surprising and strongest set of findings for the high school population was that viewing Canadian TV was

moderately associated with a *negative* attitude toward Canada. This relationship strengthens slightly with the inclusion of controls. Although the adult data contained a relationship in the same direction, it was not nearly as strong. This negative

attitude seems to be directly attributable to increased knowledge about Canada. High school viewers of Canadian TV when asked to name issues facing Canada readily listed items such as Quebec, inflation, and Margaret Trudeau; viewers of U.S. TV, by contrast, often said they knew of no problems in Canada. One obtained the impression that they viewed Canada as a big, beautiful fishing and hunting reserve.

There was little difference between Canadian and U.S. TV viewers in the issues picked as those most important for the U.S., but a substantial difference in the issues picked as those most important facing Canada and most important facing the U.S. and Canada together. These remained relatively stable with the addition of controls. These findings again support the agenda-setting hypothesis, though only for the sending country. Other sources (e.g., school) help set the agenda regarding U.S. issues for students. The paired comparison agenda-setting data for high school students (Figure 1) are less clear than for our adult population. They indicate some differences between viewers of U.S. and Canadian TV but these are not systematic and do not provide strong support for the agenda-setting hypothesis.

While for adults, viewing Canadian TV was associated with preference for nudity over violence on TV; for high school students, viewing Canadian TV was slightly associated with preference for violence over nudity on TV. This finding needs to be understood in the context of the rural conservative nature of the population. It is probable that the exposure to nudity on Canadian TV has a shock effect on the conservative rural students, while such an effect was not felt by the less naive adults. This impression is supported by informal feedback during the interview process.

For the high school students, viewing of Canadian TV was slightly associated with opposition to government housing and socialized medicine. In both cases, the effects are weak but consistent across our measures of Canadian TV viewing, and strengthened by the control *frequency of church attendance*. This agrees with the common finding of association between

conservatism and religious activity (e.g., Johnstone, 1975).

CONCLUSIONS

The data reported here lead us to several important conclusions and suggest some questions to, and clarifications of, previous cross-national TV research. First, it is apparent that there are very real cognitive effects of watching another nation's television. The fact that so many effects were found in both samples is strong evidence of the power of cross-national TV, and TV in general. This is especially so since up to 36% of Canadian TV programming is American. The effects did not appear for both sending and receiving countries as had been expected, however. Respondents viewing Canadian TV knew a great deal more about Canada, as would be expected from the work of Lin and Burt (1975), but virtually the same amount about their own country, in contradiction to the concerns of the politicians Payne and Peake (1977) describe. My first working hypothesis is, then, only partially supported. The finding that viewers of Canadian TV knew as much about the U.S. as viewers of U.S. TV is probably the result both of U.S. programming and information on Canadian TV, and spill-over from other U.S. information sources.

The viewers of Canadian TV acquire information about the United States from the American programming and news about America on Canadian TV as well as information about Canada from the Canadian programming and Canadian news content. If Canadian TV contained totally indigenous programming and much less news about the U.S., we would expect the cultural loss to be much greater. This would, of course, be the case in countries which import most of their programming, both entertainment and news. Varis (1974) documents the extent to which such massive importation takes place.

The associations between viewing of Canadian TV and knowledge about Canada are particularly impressive because of their stability even when 24 controls for other exposure to Canada, other

media use, and viewer status were employed. The drop in the association between viewing Canadian TV and language familiarity in the high school sample when frequency of listening to Canadian radio was employed as a control actually substantiates the position that the media have cross-national effects, and tentatively extends that position to include radio. Even employing this control, however, TV maintained a substantial effect. This agrees well with Sparkes's (1977) conclusion from a similar situation that about 70% of the knowledge American viewers of Canadian TV had about Canada came from watching Canadian TV.

Second, there were mixed findings about the effect of cross-cultural TV viewing on viewer's attitudes. The data did not indicate that viewers of foreign TV, when compared with viewers of local TV, had a less favorable attitude toward their own country or toward groups in that country. Indeed the only relationship of any size indicated viewers of Canadian TV were *more* favorable to American blacks. These data, therefore, do not support the concern of Nordenstreng and Varis (1973), Beattie (1967), and Schiller (1974) that foreign TV viewing would be damaging to the viewer's evaluation of his own country or parts of it. It should be remembered, however, that the messages of Canadian TV are not specifically designed to create negative images of the U.S. Carefully designed messages of this type, as a part of regular television programming, might have an effect. The data do lend support to the position of Tsai (1970) that these global attitudes about one's own country have multiple sources and are rather resistant to manipulation by TV.

In contrast to these findings about feelings for one's own country, viewers of Canadian TV reported substantially *less* favorable attitudes toward Canada than viewers of U.S. TV. This is directly opposed to our second working hypothesis and the findings of Payne and Peake, (1977) and Caron (1977). The contradiction is understandable in the context of the programming. In this study, viewers of Canadian TV saw the regular CBC programming which contained information about the problems facing Canadian soci-

ety as well as favorable information. The broadcasts from which Payne and Peake and Caron measured effects, in contrast, were more controlled, and presented a more positive image of the sending country. It appears, then, that in the absence of other information about the sending country, TV may shape people's attitudes about it. If the information is controlled to create a positive image, the positive effect may result. If, on the other hand, factual information containing both good and bad images is presented, it may remove the favorable illusions of the naive watchers and create a less favorable attitude toward the sending country. The fact that these effects are stronger for high school students than for adults indicates further the importance of naiveness in the operation of this effect. The less naive adults were less influenced by TV.

The findings that viewing Canadian or U.S. TV made no difference in attitude toward French- as compared with English-speaking Canadians indicate that if English language CBC is anti-French-Canadian, then U.S. TV is similarly biased. In contrast, adult but not high school viewers of U.S. TV do have slightly more negative attitudes toward U.S. blacks than viewers of Canadian TV. This gives some support to the concern of racial bias on U.S. TV.

Our data lend only mixed support for the agenda-setting hypothesis. Viewing of Canadian as opposed to U.S. TV was clearly related to the measures of the most important issue for Canada and for the United States and Canada together. On the other hand, the findings about the relationship between viewing Canadian TV and selection of the most important issue facing the United States were inconclusive. The open-ended questions showed only a slight and inconsistent relationship. Our paired comparison items, by contrast, indicated a clear difference in agenda-ranking for adult but not high school viewers of Canadian TV. These data fit nicely with those from the other attitudinal indicators. In the absence of other information sources (about the sending country) TV has a substantial role in setting agenda. When there are many other sources of information (about the receiv-

ing country), the role is smaller and less consistent.

Given that the media content of the two countries is similar, and that effects were found where they were most reasonably predicted by the agenda-setting theorists, these data are supportive of the work of McCombs (1976a), Siune and Borre (1975) and the host of other agenda-setting theorists. In situations where vastly differing agendas exist, for example TV signals from East and West Berlin, substantial effects on viewers might be expected.

Our measures of attitude toward government-supported housing and socialized medicine, once the controls have been employed, also fit nicely into the pattern of the other attitude measures. There appears to be no relationship between viewing Canadian TV and attitude toward these issues for adults. For high school students, viewing Canadian TV was associated with a negative attitude toward these programs. In the absence of other information sources, exposure to Canadian TV which contains negative as well as positive information about these projects predisposes the high school students against them. The less naive adult population is unaffected by the TV programming.

My third working hypothesis was that cognitive effects would be stronger than attitudinal effects. This hypothesis received clear support from these data, and agree with the notion of Tsai (1970) that it is easier to change information than basic attitudinal dispositions. It is easier to learn Canadian language and issues than to adopt Canadian attitudes. The language measures showed especially clear changes perhaps because our measures only indicated ability to identify terms, not the degree to which these terms were used by respondents in their daily speech. It is obvious that awareness and use are different issues and the former more easily acquired than the latter.

Although this paper is about the effects of television, the conclusions drawn from this research have broader implications. Read (1976), Varis (1974), Tunstall (1977), and others have documented the degree to which many American media, from the international wire services, to the comic

books, to the popular rock recordings have penetrated the world market. Findings similar to those of this study might be expected when examining effects of these other media cross-nationally. Since separate cross-cultural literature exist for the effects of a number of them, a review of each of these is beyond the scope of this paper. Examples of recent treatments of the cross-cultural effects of telephone, radio, wire service, and newspapers are provided respectively by Dicks (1977), Hudson (1977), and Read (1976). At the same time some differences are expected across even such related media as radio and television, as Wilson (1974) and Katz and Gurevitch (1976) demonstrate.

This paper is about developmentally similar countries. A synthesis of the massive literature on the role of mass media in developing countries is not intended nor within the scope of this paper. Numerous books reviewing that literature are available (e.g., Rogers, 1962; Rogers and Shoemaker, 1971; Schramm, 1964; Lerner and Schramm, 1967). It is apparent that there are many similarities between the role of TV in developmentally dissimilar countries and developmentally similar countries. Some of these are the result of technological differences. Adoption of change, particularly cognitive and behavioral change, should be more pronounced in developing countries.

Differences should also appear in each country due to the different amounts of foreign TV programming on local TV. Canadian TV is 36% U.S. programming, but many countries import more than 60% of their TV programming (Varis, 1974). Not all of that programming is U.S. and programming from other cultures may contain messages which contradict those of U.S. TV. Thus, the mix of imported TV is also an important factor. Whether the TV programming is aimed at transferring specific information, attitudes, and behavior, or is simply filling entertainment time may also affect the strength of the effect and its direction. This is made apparent by the contrasting findings of this study and those of Caron (1977).

In addition, the amount of cultural difference between the sending and receiving countries should affect the exact nature of

the transfer of attitudes, information, and behavior. Where language and values differ greatly the foreign material may be less likely to find acceptance. For all these reasons some differences between the findings reported here and those related to the effect of TV in less developed countries would not be unexpected.

Finally, this paper has not drawn comparisons with the uses and gratification literature (e.g., Katz et al., 1973; Katz and Gurevitch, 1976; Blumler and Katz, 1974). It is entirely probable that the more active audience they envision could, by having differing need distributions, affect relationships differently in different countries.

In summary, in these data viewing Canadian TV was associated with a number of attitudes and with information about both Canada and the U.S. Generally, cognitive are stronger than attitudinal effects, as I proposed, and contrary to my expectations, those toward sending, stronger than receiving country. Effects were found when we used measures at both the general and specific levels. The addition of controls for 24 media, status, and Canadian contact variables in most cases does not appreciably alter the findings.

While these conclusions are drawn from cross-sectional data and, therefore, statements about causation are limited, the study makes several advances over previous research. First, viewing of foreign TV is geographically determined, not the result of self-selection. It would be hard to argue that someone's preference for Canadian TV caused them to move to that location. Second, exposure to the nationality of TV signal has been long-rather than short-term as in many experiments. Third, the role of nonmedia exposure to Canada has been carefully controlled and evaluated. Fourth, the possible realm of effects has been systemized and several different dimensions measured so their relative strength of association can be compared.

Although the findings are important, they can not simply be extended to all mass media or all cross-national settings particularly to settings where large technological differences exist between the sending and receiving country. This

study does provide, however, comparative data about the effects of television on a number of cognitive and attitudinal dimensions at both the general and specific levels in one cross-national setting. It suggests a useful format for other research, provides a base for future comparison, and exposes areas where better or more varied measures would be useful.

REFERENCES

- Almaney, Adnan
1972 "Governments' resistance to free flow of international communication." *Journal of Communication* 22:77-88.
- Arnove, Robert F.
1975 "Sociopolitical implication of educational television." *Journal of Communication* 25:144-56.
- Beattie, E.
1967 "In Canada's centennial year: U.S. mass media influence probed." *Journalism Quarterly* 44:667-72.
- Blumler, Jay and Elihu Katz (eds.)
1974 *The Uses of Mass Communications: Current Perspectives on Gratifications Research*. Beverly Hills: Sage.
- Caron, Andre H.
1977 *The Effects of Television on Inuit Children's Cultural 'Images' of Their Own and Other Groups*. Montreal: Report to the Canadian Radio Television Commission.
- Contreras, Eduardo, James Larson, John Mayo, and Peter Spain
1976 *Cross-Cultural Broadcasting*. New York: Unesco.
- Davey, Keith
1970 *Report of the Special Senate Committee on Mass Media*. Ottawa: Queen's Printer.
1972 "The Davey report in retrospect." Pp. 171-81 in B. Singer (ed.), *Communication in Canadian Society*. Toronto: Copp Clark.
- Dicks, Dennis
1977 "From dog sled to dial phone: a cultural gap?" *Journal of Communication*. 27:120-9.
- Dunton, Davidson A. and Andre Laurendeau
1965 *A Preliminary Report of the Royal Commission on Bilingualism and Biculturalism*. Ottawa: Queen's Printer.
- Elkin, F.
1972 "Communications media and identity formation in Canada." Pp. 216-30 in B. Singer (ed.), *Communication in Canadian Society*. Toronto: Copp Clark.
- Guilford, J. P.
1954 *Psychometric Methods*. New York: McGraw-Hill.
- Horowitz, Andrew R.
1975 "The global bonanza of American TV." *More* 5:6-22.
- Hudson, Heather
1977 "The role of radio in the Canadian North." *Journal of Communication* 27:130-9.

- Johnstone, Ronald
1975 *Religion and Society in Interaction: The Sociology of Religion*. Englewood Cliffs: Prentice Hall.
- Katz, Elihu, Jay Blumler and Michael Gurevitch
1973 "Uses and gratifications research." *Public Opinion Quarterly* 37:509-23.
- Katz, Elihu and Michael Gurevitch
1976 *The Secularization of Leisure: Culture and Communication in Israel*. London: Faber and Faber.
- Kim, J.
1975 "Multivariate analysis of ordinal variables." *American Journal of Sociology* 81:261-98.
- Labovitz, Sanford
1967 "Some observations on measurement and statistics." *Social Forces* 46:151-60.
1970 "The assignment of numbers to rank order categories." *American Sociological Review* 35:515-20.
- Lerner, Daniel and Wilbur Schramm
1967 *Communications and Change in Developing Countries*. Honolulu: Eastwest Center Press.
- Lin, N. and R. Burt
1975 "Differential effects of information channels in the process of innovation diffusion." *Social Forces* 54:256-74.
- McCombs, Maxwell
1976a "Elaborating the agenda-setting influence of mass communication." Prepared for the Bulletin of the Institute for Communication Research, Keio University, Tokyo.
1976b "Empirical findings on the impact of international communication." Paper presented at a conference on Cultural Imperialism, Eastwest Center, University of Hawaii.
- Madigan, R. G. and Jack Peterson
1977 "Television on the Bering Strait." *Journal of Communication* 27:183-92.
- Nie, Norman H., C. Hadlai Hyll, Jean G. Jenkins, Karin Steinbrenner and Dale H. Bent
1975 *SPSS: Statistical Package for the Social Sciences*: 2nd ed. New York: McGraw Hill.
- Nordenstreng, Kaarle and Tapio Varis
1973 "The nonhomogeneity of the national state and the international flow of communication." Pp. 393-412 in George Gerbner, Larry Gross, and William Melody (eds.), *Communication Technology and Social Policy*. New York: Wiley.
- Payne, David E.
1977 "The effects of Canadian TV on the attitudes, values, and perceptions of viewers in Minnesota." Paper presented at the International Communications Association Meeting, Berlin.
1978 "U.S. TV in Iceland: a synthesis of studies." Paper presented at the International Communications Association meeting, Chicago.
- Payne, David E. and C. A. Peake
1977 "Cultural diffusion: the role of U.S. TV in Iceland." *Journalism Quarterly* 54:523-31.
- Peers, Frank W.
1972 "Broadcasting and national unity." Pp. 202-15 in B. Singer (ed.), *Communication in Canadian Society*. Toronto: Copp Clark.
- Read, W. H.
1976 *America's Mass Media Merchants*. Baltimore: Johns Hopkins University Press.
- Rivers, William
1971 "The Negro and the news: a case study." Pp. 151-69 in W. Schramm and D. Roberts (eds.), *The Process and Effects of Mass Communication*. Urbana: University of Illinois Press.
- Robinson, John P.
1972 "Mass communication and information diffusion." Pp. 71-93 in Gerald F. Kline and Philip Tichenor (eds.), *Current Perspectives in Mass Communications Research*. Beverly Hills: Sage.
- Rogers, Everett
1962 *Diffusion of Innovations*. New York: Free Press.
- Rogers, Everett and Floyd Shoemaker
1971 *Communication of Innovations*. 2nd ed. New York: Free Press.
- Schiller, Herbert
1973 "Authentic national development versus the free flow of information and the new communications technology." Pp. 483-503 in George Gerbner, Larry Gross, and William Melody (eds.), *Communications Technology and Social Policy*. New York: Wiley.
1974 "Freedom from the 'free flow'." *Journal of Communication* 24:110-7.
- Schramm, W.
1964 *Mass Media and National Development: The Role of Information on Developing Countries*. Stanford: Stanford University Press.
- Schulke, Beverly
1977 "Effects of television on rural Minnesota viewer attitudes toward capital punishment." M.A. thesis, Department of Sociology, University of North Dakota.
- Singer, B.
1971 "Violence, protest and war in television news: the U.S. and Canada compared." *Public Opinion Quarterly* 34:234-41.
- Siune, Kare and Ole Borre
1975 "Setting the agenda for a Danish election." *Journal of Communication* 25:65-73.
- Skipper, James K., Jr.
1975 "Musical tastes of Canadian and American college students: an examination of the massification and Americanization theses." *Canadian Journal of Sociology* 1:49-59.
- Sparkes, Vernone M.
1977 "TV across the Canadian border: does it matter?" *Journal of Communication* 27:40-7.
- Strode, Beverly
1978 "Cultural diffusion: the role of Canadian TV in the U.S." Unpublished paper. Department of Sociology, University of North Dakota, Grand Forks.
- Tsai, Michel Kuan
1970 "Some effects of American television pro-

- grams on children in Formosa." *Journal of Broadcasting* 14:229-38.
- Tunstall, Jeremy
1977 *The Media Are American: Anglo-American Media in the World*. New York: Columbia University Press.
- Varis, Tapio
1974 "Global traffic in television." *Journal of Communication* 24:102-9.
- Wilson, Edward
1974 "The effect of medium on loss of information." *Journalism Quarterly* 51:111-5.

BLACK POLITICAL PARTICIPATION IN THE UNITED STATES: SOME RECENT EVIDENCE*

NICHOLAS L. DANIGELIS

University of Vermont

American Sociological Review 1978, Vol. 43 (October):756-771

A recently presented political climate theory of black political participation is assessed against indirect effects, isolation and ethnic community arguments on black political activity. Multivariate trend analysis of the University of Michigan presidential election year national surveys from 1952 through 1972 support a modified political climate theory as being most appropriate to understanding differences in voting and vote-related activity levels between white and black Americans in recent years.

Social and political research has produced a plethora of apparently contradictory research findings which generate quite disparate views on the actual political involvement of blacks in American life. Recently, a composite theory based on selected assumptions from three competing theories of black participation, called a *political climate* theory of participation (Danigelis, 1977), has been offered as a means of resolving contradictions. Unfortunately, neither the political climate theory nor any of the explanations it purportedly replaces has been subjected to empirical test. Focusing on conventional political behavior, the present paper contrasts white and black political activity during the period 1952-1972 with multivariate trend analysis of survey data from six national studies.

*Address all communications to: Nicholas L. Danigelis; Department of Sociology; University of Vermont; Burlington, VT 05401.

Much appreciation is due the University of Michigan Survey Research Center which kindly allowed use of its data. Special thanks for helpful advice and criticism are given Elton Jackson, Shel Stryker, Gerry Marwell and the two anonymous ASR referees who read my manuscript. For diligent research help I owe Pete Yeager a deep debt of gratitude. Responsibility for data manipulation and interpretation rests solely with me.

THE ARGUMENTS

Three Popular Explanations of Black Political Participation

Used to explain both political and social participation levels among American blacks are three theories which have been generated by post-World War II survey research. The first argues that black political involvement is explained on the basis of nonrace factors like socioeconomic status:

We know that SES is positively related to conventional and unconventional political participation. We also know that blacks are more likely to be found in the lower strata of society. The question therefore arises, is it the low SES of Negroes that impedes their participation, or is it their blackness? This question has been investigated rather thoroughly and we can confidently state that the somewhat lower participation among blacks is a reflection of their lower SES, and not of any racial characteristics. (Milbrath and Goel, 1977:120)

While plausible, this *indirect effects* explanation is incomplete at best, because black Americans historically also have been disproportionately among the young (Olsen, 1970:683), in the rural South (Pinkney, 1975:47), in nonunion

families (see Simpson and Yinger, 1972:331-5, for a discussion of factors affecting this difference),¹ and in groups for whom participation rates are low (see Lane, 1959:48-9; Flanigan, 1972:22-3; and Milbrath and Goel, 1977:86ff, for summary discussions). Therefore, an inclusive indirect effects theory explains the disparity in black and white political participation rates via membership variables which are correlated jointly with race and levels of political behavior. Implied here is the assumption that, if whites were to be found in the same proportion as blacks in these various subgroups, then all political participation rate differences between blacks and whites would disappear (Olsen, 1970:682, discusses this assumption with respect to socioeconomic status).

To the question of whether it is SES or race which impedes black political participation the *isolation* theory (Orum, 1966:34) answers "both." In addition to indirect effects, this explanation argues a direct race effect based on Lipset's (1960:112) discussion of lower-class isolation. In the first place, structural barriers such as threats of violence and disfranchisement techniques are considered direct obstacles to black voting and other modes of political expression. As a result of such discrimination, blacks are said to experience both social alienation or anomie and political apathy (Orum, 1966:33), which in turn isolate them from the mainstream of society's major social, economic and political activities and, therefore, prevent them from building up the competence necessary to deal with the complexities of political life. Therefore, if the isolation theory is correct, one expects to find that, once controls for socioeconomic status, age, region and the like are made, blacks on the average should be more alienated from and less interested in politics and, therefore, less politically active than whites.

Based on multivariate analyses of racial

effects on political activity in selected urban areas (see e.g., Williams et al., 1973; Antunes and Gaitz, 1975) and consistent with several political profiles of blacks in particular geographical areas (see e.g., Gosnell, 1935; Matthews and Prothro, 1966), the *ethnic community* theory of political participation (Olsen, 1970) predicts high levels of political activity among black Americans in direct opposition to the predictions of the isolation theory. Ironically, the initial causal force in the ethnic community theory is the same as that in the isolation theory: structural barriers to black political participation. The result of these barriers, however, is not apathy or alienation but rather a sense of community and social cohesion which produces among blacks a norm of activism to which members of the newly created ethnic community are expected to conform (Olsen, 1970:684). All other things being equal (i.e., with the necessary statistical controls made), the ethnic community theory argues that discrimination produces not lower but higher levels of political efficacy and interest and, therefore, of political participation among blacks than among whites.

The Political Climate Theory of Black Political Participation

Recently (Danigelis, 1977) a composite theory combining central assumptions of the indirect effects, isolation and ethnic community arguments has been proposed. This political climate theory of black political behavior maintains that different areas and time periods in this country's history can be characterized by political climates intolerant or supportive of or indifferent to black political participation² and that, as a consequence, blacks have faced differing levels of political discrimination historically. The type of political climate, therefore, is the key to understanding the political profile of black Americans.

... [T]he political discrimination in the isolation theory is different from that in the ethnic

¹ Clearly the proportion of blacks in these groups is changing (e.g., note large migration of rural Southern blacks to urban areas in the North during the 1950s and 1960s); the only point being made is that at particular times blacks have been disproportionately found in groups with low political participation rates.

² The argument is more detailed than the summary presented here; for a more comprehensive discussion the reader should consult Danigelis (1977).

community theory due to the different political climates involved.

On the one hand, an intolerant political climate represents nearly complete restriction by law of black voting and almost unanimous white support of such restrictions. . . . Isolation, therefore, requires complete political discrimination which, in turn, derives from an intolerant political climate.

On the other hand, a supportive political climate means less than complete restriction of black political participation, because, while the dominant theme is discrimination, there is no unified white opposition to black political involvement. This incomplete political discrimination produces the milieu in which ethnic community can develop. (Danigelis, 1977:40)

The indirect effects argument is then offered to explain any white-black participation differences when the political climate is neutral or ambiguous toward black participation (Danigelis, 1977:Figure 2). While the arguments are more detailed than this brief discussion implies, the thrust of the political climate theory is that white-black differences in political participation are variable over time and place according to differences in political climate as measured by level of political discrimination.

Previous Evidence

Recent research has produced support for the indirect effects theory (e.g., Clemente and Sauer, 1975), the isolation theory (Daniel, 1969; Olsen, 1970:693-5) and the ethnic community theory (Olsen, 1970; Williams et al., 1973; 1975). Evidence supporting the political climate theory comes from two sources. First is the indisputable fact that the above cited sources as well as many others portray apparently conflicting pictures of white-black political participation differences. Such *prima facie* evidence for a nonstatic view of black politics is buttressed by over-time comparisons of registration rates between whites and blacks (Simpson and Yinger, 1972:393; Campbell and Feagin, 1975:134) and by both time and regional differences in voting rate comparisons between blacks and whites (Orum, 1966:44; Campbell and Feagin, 1975:136).

Unfortunately, as can be seen from the

above, the evidence relevant to the various arguments on black political participation have been characterized by differences not only of findings but also of time, location and measures of political participation. The need both to control these variable factors and to provide a reasonable test of each theory suggests several different methodologies (see Danigelis, 1977, for a discussion of some). One particularly valuable method is a trend analysis of national survey data for which data now exist in a fairly uniform format for a number of years since the early 1950s. Data from six studies conducted by the University of Michigan's Survey Research Center will be used to test the four theories of black political participation discussed above.

METHODOLOGY

Data

The University of Michigan's six election year studies between 1952 and 1972 have utilized similar sampling procedures and asked either identical or similar questions,³ providing many obvious advantages for multivariate analysis. Most important for this study, the data cover a time characterized by periods of substantial political, social and economic change for many segments of the black American populace as well as by periods of relative stability. Therefore, comparisons can be made between time periods to test both the durability of the three time-bound theories of black participation and the expected changes in black-white comparisons predicted by the political climate theory.

The data restrict the analysis in two ways, however: first, confinement to those variables which the original investigators deemed advisable to include means that it will be impossible to examine several different kinds and levels of political activity because questions are limited to voting behavior and participation in selected political spheres with re-

³ The reader is referred to the published codebooks for each study; included in each are descriptions of sampling techniques and the exact text and all response categories for each question.

spect to national elections only or with no specific referent. Therefore, the extent to which local political behavior differences do or do not exist cannot be gauged. A second problem stemming from using other people's data is that response categories (and, to a minor extent, wording of questions) were changed from study to study. Such differences among studies, however, are relatively minor and will not hinder statistical comparisons between studies. (See Glenn, 1970, for a discussion of trend analysis problems in general.) Overall, the following analysis falls short of a comprehensive test of black political participation theories but represents a substantial advance over recent one-shot survey analyses.

Operationalization of Concepts

Dependent variables. Political activity has been defined as "all behavior through which people directly express their political opinions" (Matthews and Prothro, 1966:37). Because this study concentrates on legitimate political behavior, attitudes and behavior regarding political protest and violence will be ignored—even though these means of expressions generally could be included in Matthews and Prothro's definition. In operational terms, political activity will encompass registration, voting and nonvoting "legitimate" political behavior. Voter registration is measured by a direct question asking whether the respondent is registered to vote: a one (1) is given those who were registered at the time of the interview and a zero (0) to those not registered, so that the unadjusted (Table 1) and adjusted (Table 4) registration comparisons between whites and blacks showing proportion of each race registered is the same as showing average registration scores for each race. Similarly, voting is measured by a postelection question asking whether a respondent had voted (1 for yes, 0 for no), so that in Tables 2 and 5 the voting proportion differences are again the same as average differences in voting.

Finally, nonvoting political activity has been measured by summing the number of "yes" answers to questions asking whether the respondent tried to convince

others to vote for one of the parties or candidates, attended political meetings or the like, worked for one of the parties or candidates, belonged to any political clubs or organizations, and wore campaign buttons or put a campaign sticker on his/her car. A "yes" to each question produces the highest score (5) while a "no" to each question gives the lowest score (0) on the participation index. For comparability and because of a highly skewed distribution in each of the studies, participation has been dichotomized: those receiving a one (1) have participated in at least one of the above activities and those receiving a zero (0) have participated in none. Therefore, in Tables 3 and 6 proportions participating may be likened to average participation scores.

At this point, it should be clear that a variety of political activity dimensions are going to be tapped (note Milbrath and Goel's 1977:10ff, discussion of the multidimensional nature of political participation). They range from the patriotic-duty activities of registration and voting, which most white Americans take for granted and for which most blacks have had to fight, to the relatively less frequent vote-related activities like joining political clubs, going to political rallies, donating money to political candidates.

Independent variables. Race is dichotomized to distinguish between white and black respondents, with those classified *other*—a negligible number in each survey—being omitted. Other variables, assumed by a comprehensive indirect effects argument to explain the gross race-political participation correlation, include several which are measured by the Michigan surveys. One is the degree to which the individual is committed to the present political party system, regardless of party affiliation. To operationalize this concept, an index, derived from a three-part question asked of all respondents concerning their party affiliation and the degree to which they are attached to that party, gives the following categories:

- A—Strong Democrat, strong Republican (4 points);
- B—Not very strong Democrat, not very strong Republican (3 points);

- C—Independent closer to Democrats,
Independent closer to Republicans
(2 points);
- D—Independent (1 point);
- E—Apolitical (0 point).

These response categories result from collapsing categories (e.g., *Strong Democrat* was listed as a different category from *Strong Republican* in the surveys). The reasons for classifying respondents thus stem from the belief that being a strong Democrat and being a strong Republican reflect a real commitment to partisan politics and that there is nothing inherent in either party to require a distinction between strong Democrats and strong Republicans. Similarly, Independents—no matter what their background—are not as partisan as the individuals in either group A or B. Finally, of course, individuals who do not vote or “don’t mess with politics” (the Apoliticals) reflect even less of a commitment than Independents. The variable thus described will be referred to hereafter as *Strength of Party Identification*. Strength of party identification should be stronger among whites, who have more experience than blacks historically in partisan politics, and also should be correlated positively with political participation.

Other variables, whose possible roles in explaining any race-participation association have been discussed earlier, include sex (males vs. females), marital status (married vs. nonmarried), age (in years, from 18 up), union membership (no one in household a union member vs. at least one person in household belonging to a union), size of city of residence (fewer than 10,000 vs. 10,000 to 99,999 vs. 100,000 and over), education (in years of schooling from zero up), occupational prestige (in Duncan SES scores), and family income (in thousands of dollars, with income category means as estimates of respondent’s family income).

Finally, the concept of political climate is based on a crude estimate of laws affecting black political activity and white attitudes toward those laws. Where laws and feelings are highly hostile toward black participation in politics, an intolerant political climate should be found. The pre-civil rights movement South

exemplifies an intolerant political climate. A supportive political climate involves at worst incomplete white resistance to black participation and at best active white support. An example of this type of climate is the South during the height of the civil rights movement. A neutral or ambiguous political climate, where discriminatory barriers are absent or basically irrelevant to black political participation, probably is found throughout much of the North during the period between 1952 and 1972.

As suggested above, the indicators of political climate are rather crude and not without problems. One simplifying assumption is that, given the nature of the data to be analyzed, a reasonable set of indicators of political climate can be based on year of survey (1952, 1956, 1960, 1964, 1968, 1972) and region of survey (North vs. South), especially because it is over time in one region (South) where the changeover from intolerant to supportive political climate can be observed as civil rights marches and Civil Rights and Voting Rights Acts produced visible changes in political barriers affecting blacks (see Danigelis, 1977:27–41, for a discussion of the relevance of such events). In the next section, these and other assumptions are operationalized in the research hypotheses.

Another simplifying assumption is that, while registration, voting and vote-related participation reflect different dimensions of political activity, the form and magnitude of black-white differences should not vary much from one indicator to another. Therefore, each hypothesis will refer to proportions of blacks and whites who are politically active, meaning the percent who are registered to vote, who voted or who participated in at least one of the vote-related activities like joining a political club, donating money, etc. Each hypothesis, however, will be tested by separate registration, voting and participation comparisons to determine whether or not this simplifying assumption is warranted.

The Hypotheses

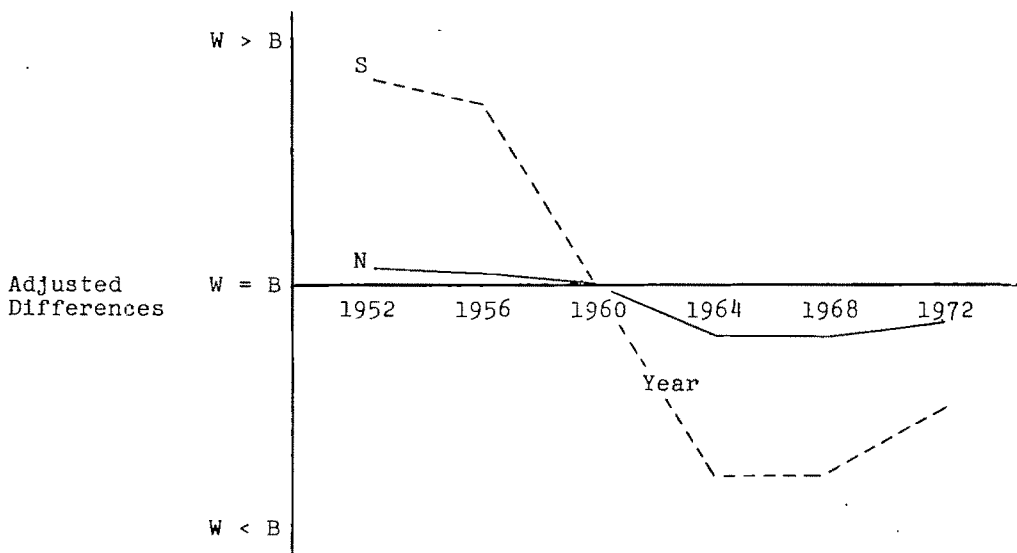
All of the theories to be tested assume that zero-order comparisons of white and

black proportion politically active should show the former to be higher, although some studies in particular areas (e.g., Gosnell, 1935:16-7; Agger et al., 1964:268; Matthews and Prothro, 1966:523) have shown the reverse to be true. Thus, hypothesis one: in unadjusted comparisons, the proportion of politically active whites *will be higher* than the proportion of politically active blacks.

The theories' predictions diverge when controls for sex, age, education and the like are instituted in a multiple regression analysis. The indirect effects argument assumes such controls will eliminate white-black proportion differences and is represented by hypothesis two: in adjusted comparisons, the proportion of politically active whites *will not differ significantly* from the proportion of politically active blacks. The isolation theory, maintaining that discrimination suppresses black political activity even after sex, age and all the other variables are controlled, gives hypothesis three: in adjusted comparisons, the proportion of politically active whites *will be higher* than the proportion of politically active blacks. In direct contradic-

tion, the ethnic community argument assumes that discrimination serves as an incentive to political activity among blacks and produces hypothesis four: in adjusted comparisons, the proportion of politically active whites *will be lower* than the proportion of politically active blacks.

The composite political climate theory argues that hypotheses two through four are specific to particular locales at specific times, depending on the level of discriminatory barriers relevant to black political life. In the North (as represented by line N in Figure 1), where the full brunt of both political discrimination and the attempts to remove it were not felt strongly, there should be small racial differences in the early period and only moderate differences in the later period, as represented by hypothesis five: adjusted white and black political activity proportion differences will follow the indirect effects prediction (no significant differences, although white proportions might be slightly higher) in 1952 through 1960 and in 1972, and the ethnic community prediction (white proportions moderately lower than black proportions) in 1964 and 1968. In the



W = Adjusted white political participation averages

B = Adjusted black political participation averages

Note: Distance from line where $W=B$ represents crude estimate of magnitude of $W-B$ differences

Figure 1. Graphic Representation of Political Climate Working Hypotheses

Table 1. Unadjusted Differences in Voter Registration Proportions between Whites and Blacks by Region and Year, 1952-1972

	1952	1956	1960	1964	1968	1972
Both Regions						
Proportion Registered						
Among Whites (R_W)	.81	.81	.83	.85	.85	.79
(Sample Size)	(1,604)	(1,592)	(1,718)	(1,379)	(2,764)	(1,199)
Among Blacks (R_B)	.36	.46	.60	.75	.77	.82
(Sample Size)	(169)	(141)	(170)	(416)	(282)	(130)
$R_W - R_B^a$	+0.45***	+0.35***	+0.23***	0.10***	+0.08**	-0.02(ns)
North Only						
Proportion Registered						
Among Whites (R_W)	.84	.82	.86	.87	.87	.81
(Sample Size)	(1,215)	(1,169)	(1,166)	(1,001)	(1,964)	(826)
Among Blacks (R_B)	.68	.67	.74	.89	.82	.80
(Sample Size)	(56)	(63)	(77)	(183)	(129)	(56)
$R_W - R_B^a$	+0.16*	+0.15*	+0.11*	-0.02(ns)	+0.05(ns)	+0.01(ns)
South Only						
Proportion Registered						
Among Whites (R_W)	.71	.77	.79	.81	.78	.75
(Sample Size)	(389)	(423)	(552)	(378)	(800)	(373)
Among Blacks (R_B)	.20	.29	.48	.64	.72	.82
(Sample Size)	(113)	(78)	(93)	(233)	(153)	(74)
$R_W - R_B^a$	+0.51***	+0.47***	+0.31***	+0.17***	+0.06(ns)	-0.08(ns)

^a A difference of proportions test assuming a conservative estimate of population standard errors was performed.

(ns) Not significant at .10 level (2-tailed test).

* Significant at .10 level (2-tailed test).

** Significant at .01 level (2-tailed test).

*** Significant at .001 level (2-tailed test).

Note: Some differences, when added to the black proportion do not equal the white proportion due to rounding.

South (see Figure 1, line S), where political discrimination was greater and the fight over it more intense, the expectation is that racial differences will be more extreme and will change more dramatically, as predicted by hypothesis six: adjusted white and black political activity proportion differences will follow the isolation prediction (white proportions substantially higher than black proportions) in 1952 and 1956, the indirect effects prediction (no significant difference) in the transition period 1960, and the ethnic community prediction (white proportions substantially lower than black proportions) in 1964 through 1972.

The hypotheses of the political climate argument are based on evidence that, following the political discrimination (especially in the South) of the 1950s, white support for and interest in full black political participation reached a peak in the mid-1960s and began to wane in the late 1960s and early 1970s as the moderate coalition rhetoric of the civil rights movement became replaced by the self-help militance of black-power rhetoric

(Downes and Burks, 1969:338; Lomax, 1971:283-300).⁴ Because of this reduced white support, it is reasonable to expect that the momentum toward full political equality by blacks may have received enough of a setback to reduce the effectiveness of those ethnic community feelings which were so prevalent among blacks at the height of the civil rights movement. Therefore, both the North and South adjusted white-black comparisons in 1972 should still be negative but should show a smaller difference than comparisons in 1964 and 1968.

RESULTS

Unadjusted Differences

Registration and voting proportion differences for each year for both regions combined (Tables 1 and 2, first row dif-

⁴ Killian (1968:130) rightly points out that much of the violent rhetoric associated with the black protest movement came from white radicals; nevertheless, white money and white liberal support was being turned off as the threats got louder and the language hotter.

Table 2. Unadjusted Differences in Voting Proportions between Whites and Blacks by Region and Year, 1952-1972

	1952	1956	1960	1964	1968	1972
Both Regions						
Proportion Voting						
Among Whites (V_w)	.78	.77	.82	.80	.77	.74
(Sample Size)	(1,453)	(1,608)	(1,657)	(1,287)	(2,474)	(2,032)
Among Blacks (V_B)	.33	.35	.53	.68	.65	.65
(Sample Size)	(157)	(145)	(152)	(382)	(243)	(221)
$V_w - V_B^a$	+0.45***	+0.41***	+0.29***	+0.12***	+0.12***	+0.09**
North Only						
Proportion Voting						
among Whites (V_w)	.84	.82	.86	.83	.81	.78
(Sample Size)	(1,122)	(1,178)	(1,130)	(932)	(1,752)	(1,403)
Among Blacks (V_B)	.71	.59	.70	.84	.71	.71
(Sample Size)	(49)	(63)	(70)	(165)	(111)	(83)
$V_w - V_B^a$	+0.12*	+0.23***	+0.16**	-0.01(ns)	+0.10*	+0.07(ns)
South Only						
Proportion Voting						
Among Whites (V_w)	.60	.62	.73	.71	.68	.63
(Sample Size)	(331)	(430)	(527)	(355)	(722)	(629)
Among Blacks (V_B)	.16	.17	.39	.56	.61	.61
(Sample Size)	(108)	(82)	(82)	(217)	(132)	(138)
$V_w - V_B^a$	+0.44***	+0.45***	+0.34***	+0.16***	+0.08*	+0.03(ns)

See Table 1 footnotes for explanations of ^a, (ns), *, **, ***.

ferences) show considerable support for hypothesis one, because five of six registration and all voting comparisons show whites to be significantly more active than blacks. Participation differences (Table 3, first row differences), however, while all positive as predicted, are significant only for the first three years. Note also that the

large proportion differences in 1952 for each of the first two tables are reduced in 1956, then again for each subsequent year, the smallest difference being in 1972. In Table 3, the participation differences are similarly highest in 1952 and lowest in 1972 but do not show the same consistent year-by-year reduction in the last three

Table 3. Unadjusted Differences in Participation Proportions between Whites and Blacks by Region and Year, 1952-1972

	1952	1956	1960	1964	1968	1972
Both Regions						
Proportion Participating						
Among Whites (P_w)	.33	.35	.41	.38	.39	.39
(Sample Size)	(1,443)	(1,603)	(1,645)	(1,280)	(2,366)	(1,942)
Among Blacks (P_B)	.14	.24	.31	.37	.36	.38
(Sample Size)	(156)	(145)	(153)	(374)	(240)	(216)
$P_w - P_B^a$	+0.19***	+0.11*	+0.10*	+0.01(ns)	+0.03(ns)	+0.01(ns)
North Only						
Proportion Participating						
Among Whites (P_w)	.34	.36	.40	.38	.40	.39
(Sample Size)	(1,115)	(1,175)	(1,123)	(925)	(1,686)	(1,331)
Among Blacks (P_B)	.10	.34	.51	.39	.35	.51
(Sample Size)	(49)	(62)	(70)	(158)	(109)	(78)
$P_w - P_B^a$	+0.24**	+0.02(ns)	-0.11*	-0.00(ns)	+0.05(ns)	-0.12*
South Only						
Proportion Participating						
Among Whites (P_w)	.28	.34	.41	.38	.38	.39
(Sample Size)	(328)	(428)	(522)	(355)	(680)	(611)
Among Blacks (P_B)	.16	.17	.13	.36	.37	.30
(Sample Size)	(107)	(83)	(83)	(216)	(131)	(138)
$P_w - P_B^a$	+0.12*	+0.17**	+0.28***	+0.02(ns)	+0.01(ns)	+0.08*

See Table 1 footnotes for explanation of ^a, (ns), *, **, ***.

years where each difference is insignificant.

A breakdown by region as well as year for each political activity comparison yields interesting divergencies (second and third row differences). In Table 1, registration comparisons in the North are similar to the general participation pattern just discussed: the first three years show significant positive differences, while the last three years show insignificant differences. In the South, the trend is almost linear, as the large positive difference (+.51) in 1952 decreases steadily until it becomes negative in 1972 (-.08) (not significant, however). In Table 2, there is much fluctuation in the North over time as whites are participants to a significantly greater extent in four years while, in 1964 and 1972, the differences are nonsignificant. In the South, the trend seen in Table 1 is evident here as the larger positive difference in 1952 is reduced year after year until 1972 where a nonsignificant difference is found. In Table 3, in the North, participation differences go from significantly positive to significantly negative between 1952 and 1960, then are nonsignificant for two years and again negative and significant in 1972. In the South, all differences are positive but those in the first three years are all significant and larger than those in the last three years (two of the latter being nonsignificant).

Altogether, the 36 region- and year-specific registration, voting and participation comparisons yield 21 significant positive differences supportive of hypothesis one, two significant negative differences, and 13 comparisons with no differences. The most supportive data are for the years 1952, 1956 and 1960 where 16 out of 18 differences are positive and significant; the years 1964, 1968 and 1972 yielded only five positive significant differences out of 18 possible. Therefore, while total sample white-black comparisons for each year support hypothesis one consistently in terms of registration and voting, the comparable participation comparisons and the region-specific comparisons for all three dependent variables are not nearly as supportive; they show some apparent interaction among region, year and type of political activity.

Adjusted Differences

Once sex, age, education and the like are controlled, the direct effect of race becomes much clearer. The underlying assumption of the indirect effects argument, and of all the other theories, is that correlates of both race and political activity put blacks at a disadvantage to whites in the political arena. Blacks are represented disproportionately among very young, the poorly educated, low income families, etc., and it is in these groups that low voter turnout and few participants are to be found. Therefore, multiple regression analysis should reduce positive zero-order differences and increase negative zero-order differences.

Registration comparison. In fact, the differences between unadjusted registration comparisons in Tables 1 and 4 are, with only one exception, in the anticipated directions: the differences for each year become approximately halved (1952 through 1968) or become even stronger if originally negative (1972). When both region and year are held constant simultaneously, the same pattern obtains with the exceptions of 1968 and 1972 in the North and 1968 in the South (although all but the 1972 North comparison show a decrease in magnitude of the positive coefficients, as predicted).

Hypothesis two is not supported by Table 4 because only seven differences are negligible (between $\pm .05$ and 0) and nonsignificant ($p > .10$), while only one is nonnegligible but still nonsignificant. Therefore, even with statistical significance the governing criterion, out of the six year-specific and 12 region- and year-specific comparisons (18 in all), only eight (less than half) show no difference between white and black registration profiles once relevant controls have been made.

Hypothesis three is not consistent with the data either, because only seven (slightly more than one-third) registration differences are positive and significant at the .10 level. Finding even less support, hypothesis four is consistent with only the North-only comparison for 1964 and the combined and South-only comparisons for 1972.

Table 4. Adjusted Differences in Registration Proportions between Whites and Blacks by Region and Year, 1952-1972

	1952	1956	1960	1964	1968	1972
Both Regions						
$R_W - R_B^a$	+0.26***	+0.19***	+0.10**	+0.02(ns)	+0.02(ns)	-0.08*
Standard Error	.03	.04	.03	.02	.02	.04
(Sample Size)	(1,727)	(1,703)	(1,839)	(4,488)	(3,022)	(1,297)
North Only						
$R_W - R_B^a$	+0.08(ns)	+0.04(ns)	+0.05(ns)	-0.06*	+0.03(ns)	+0.03(ns)
Standard Error	.05	.05	.04	.03	.03	.05
(Sample Size)	(1,240)	(1,211)	(1,211)	(3,143)	(2,077)	(863)
South Only						
$R_W - R_B^a$	+0.34***	+0.27***	+0.09*	+0.06*	+0.01(ns)	-0.18***
Standard Error	.05	.06	.05	.03	.04	.06
(Sample Size)	(487)	(492)	(628)	(1,345)	(945)	(434)

^a Each adjusted proportion difference is the race dummy regress on coefficient (white=1, black=0) in a multiple regression analysis for which independent variables are described in the text.

(ns) Not significant at .10 level (2-tailed F test).

* Significant at .10 level (2-tailed F test).

** Significant at .01 level (2-tailed F test).

*** Significant at .001 level (2-tailed F test).

If only the 12 region- and year-specific comparisons are considered the picture remains essentially the same: out of 12 registration comparisons, six are non-significant, four are positive and two negative.

Concentrating on these 12 comparisons also allows a test of hypotheses five and six which predict separate regional trend

changes in the sign of the adjusted registration differences. Comparison of Figure 1's projected proportion political participation differences and Figure 2's actual proportion registration differences partially supports the political climate theory's hypotheses. In the North, differences are positive and nonsignificant in 1952 through 1960, negative and signifi-

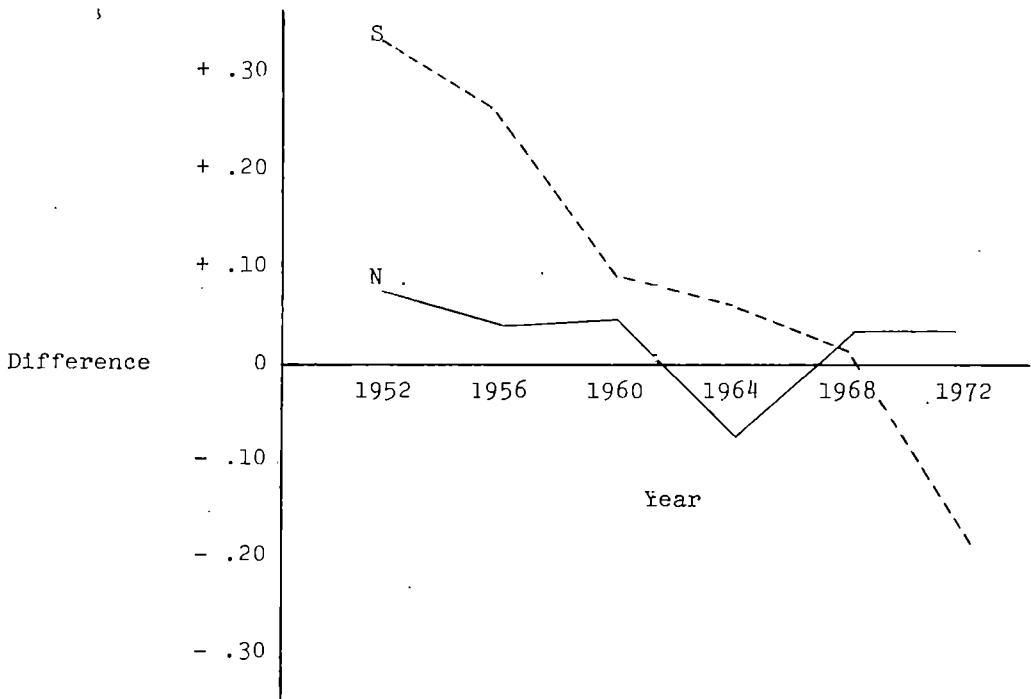


Figure 2. Adjusted Differences in Proportion Registered Plotted by Sign and Magnitude, by Region and by Year

cant in 1964 and insignificant again (although positive instead of negative) in 1972; unexpected is the positive (but non-significant) difference in 1968. In the South, as predicted by hypothesis six, the trend over time is for large positive differences to be reduced and eventually to change to negative differences; unexpectedly, however, there is no leveling off, suggesting a negative linear rather than curvilinear relationship between time and registration differences in the South. Also, in the South the transition period between positive and negative differences occurs not in 1960 as expected but rather in 1968.

Voting comparison. The differences between unadjusted and adjusted comparisons found for registration also obtain for voting (Tables 2 and 5). Large positive coefficients are reduced in size by about half, marginal positive coefficients become negative and negative ones stronger, the exceptions being in the North in 1960, where the reduction in size of the coefficient is marginal, and in 1972, where the coefficient remains the same size. Generally, the most dramatic changes are found in the South.

As in the registration results, hypothesis two receives marginal support because only six coefficients are between $\pm .05$ and 0 in magnitude and statistically nonsignificant (another two are nonnegligible but also not significant) out of the 18 coefficients in Table 5. Nor is hypothesis three any better served by the data, because again only eight are both positive and significant at the .10 level. Finally, hypothesis four continues to receive the

least amount of support, as only two coefficients are negative and statistically significant. An examination of the 12 region- and year-specific coefficients does not alter the picture: five are nonsignificant, five positive and two negative.

If the adjusted coefficients for the North are plotted (Figure 3), it is clear the only expected results are the marginal positive coefficient in 1952 and the moderately negative coefficient in 1964; unexpected are the strong positive coefficients in 1956 and 1960 and the weak positive coefficients in 1968 and 1972. The curve for the South follows the one for the South in Figure 2. Voting differences are positive and large in the early years, are reduced and become negative by 1972. In fact, if the curves in Figures 2 and 3 are compared, the signs of the registration and voting coefficients are identical when specified by region and year. The only major difference, in fact, is that the two voting coefficients for the North for 1956 and 1960 are large and significant, while the respective registration coefficients are small and nonsignificant, although all are positive. Therefore, hypotheses five and six receive approximately the same degree of support from the voting as from registration results.

Participation comparison. Here, as with the registration and voting differences, the tendency is for unadjusted positive differences to become less positive or to change signs and for unadjusted negative differences to become stronger (Tables 3 and 6). Two minor exceptions are in the North, where in 1968 and 1972

Table 5. Adjusted Differences in Voting Proportions between Whites and Blacks by Region and Year, 1952-1972

	1952	1956	1960	1964	1968	1972
Both Regions						
$V_w - V_B^a$	+0.20***	+0.21***	+0.14***	-0.01(ns)	+0.03(ns)	-0.03(ns)
Standard Error	.04	.04	.03	.02	.03	.03
(Sample Size)	(1,570)	(1,723)	(1,746)	(4,192)	(2,695)	(2,202)
North Only						
$V_w - V_B^a$	+0.04(ns)	+0.15**	+0.13**	-0.06*	+0.06(ns)	+0.07(ns)
Standard Error	.06	.05	.04	.03	.04	.05
(Sample Size)	(1,144)	(1,220)	(1,158)	(2,931)	(1,847)	(1,459)
South Only						
$V_w - V_B^a$	+0.26***	+0.19**	+0.13*	+0.01(ns)	+0.01(ns)	-0.14**
Standard Error	.05	.06	.05	.04	.05	.05
(Sample Size)	(426)	(503)	(588)	(1,261)	(848)	(743)

See Table 4 footnotes for explanations of ^a, (ns), *, **, ***.

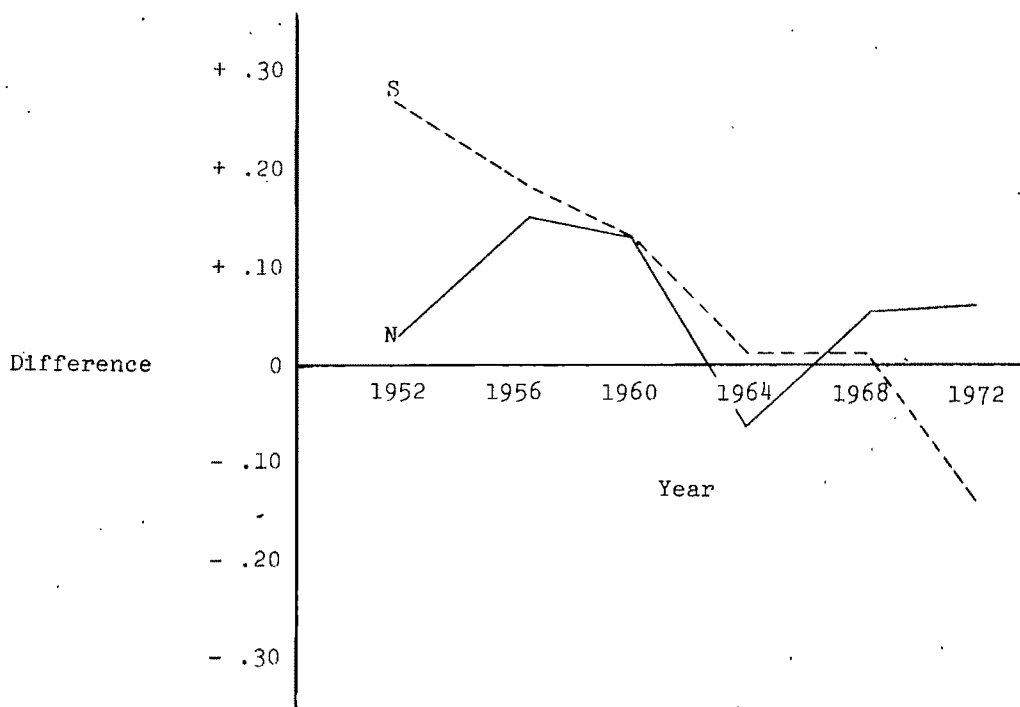


Figure 3. Adjusted Differences in Proportion Voting Plotted by Sign and Magnitude, by Region and by Year

positive differences do not change appreciably (although in 1972 the unadjusted difference is significant and the adjusted difference insignificant).

Hypothesis two receives better support from Table 6 than from either Table 4 or 5, since 13 of 18 participation comparisons are statistically nonsignificant ($p > .10$). Hypothesis three receives far less support here than in the previous comparisons, because only two differences (1952 for both regions and in the North) are positive

and significant. Hypothesis four receives about the same support as in the previous comparisons, since only three differences are negative and significant.

Out of the 12 comparisons controlling for region and year, nine (four North and five South) show no significant difference between white and black participants. The 1952 comparisons in the North show whites to be more politically active and the 1960 North and 1964 South comparisons show blacks to be more active.

Table 6. Adjusted Differences in Participation Proportions between Whites and Blacks by Region and Year, 1952-1972

	1952	1956	1960	1964	1968	1972
Both Regions						
$P_w - P_b^a$	+0.07*	-0.03(ns)	-0.06(ns)	-0.08**	-0.00(ns)	-0.04(ns)
Standard Error	.04	.04	.04	.03	.03	.04
(Sample Size)	(1,559)	(1,715)	(1,735)	(4,172)	(2,578)	(2,106)
North Only						
$P_w - P_b^a$	+0.18*	-0.07(ns)	-0.22***	-0.04(ns)	+0.04(ns)	-0.09(ns)
Standard Error	.07	.06	.06	.04	.05	.06
(Sample Size)	(1,137)	(1,215)	(1,151)	(2,912)	(1,775)	(1,381)
South Only						
$P_w - P_b^a$	+0.02(ns)	+0.03(ns)	+0.05(ns)	-0.08*	-0.07(ns)	-0.03(ns)
Standard Error	.05	.06	.06	.04	.05	.05
(Sample Size)	(422)	(500)	(584)	(1,260)	(803)	(725)

See Table 4 footnotes for explanations of ^a, (ns), *, **, ***.

The over-time changes of these 12 comparisons are plotted in Figure 4. For the North the results do not conform to the predictions of hypothesis five. Initially the differences are positive, change to negative, then become positive and finally negative once again. Most unexpected is the strong negative coefficient that appears as early as 1960 and is followed by a pair of negligible positive coefficients in 1964 and 1968 when the coefficients should have been negative. In the South there is a trend resembling the prediction of hypothesis six, with all three of the early differences positive and the last three negative; however, most (five of six) differences are insignificant, showing that what trend there may be is slight at best. In the participation comparisons, therefore, the indirect effects argument receives the most support among the competing theories for the North and South, although the expected change from early small positive coefficients to a moderate negative one in 1964 does obtain in the South.

DISCUSSION

Trend analysis based on the twenty-year period between 1952 and 1972 appears to indicate that two theories of racial differences in political participation—political climate and indirect effects—are appropriate to a discussion of black and white political activity comparisons in the post-World War II United States. Specifically, registration and voting comparisons support the political climate theory of participation, while comparisons of political activity other than registration and voting are consistent with the indirect effects argument and, in the South, both the indirect effects and political climate theories. The reasons probably are due to the parameters surrounding the applicability of political climate to the relationship between race and political participation.

Political climate. Before explanations for the unexpected results are attempted, those predictions upheld by the data should be emphasized. First, and consis-

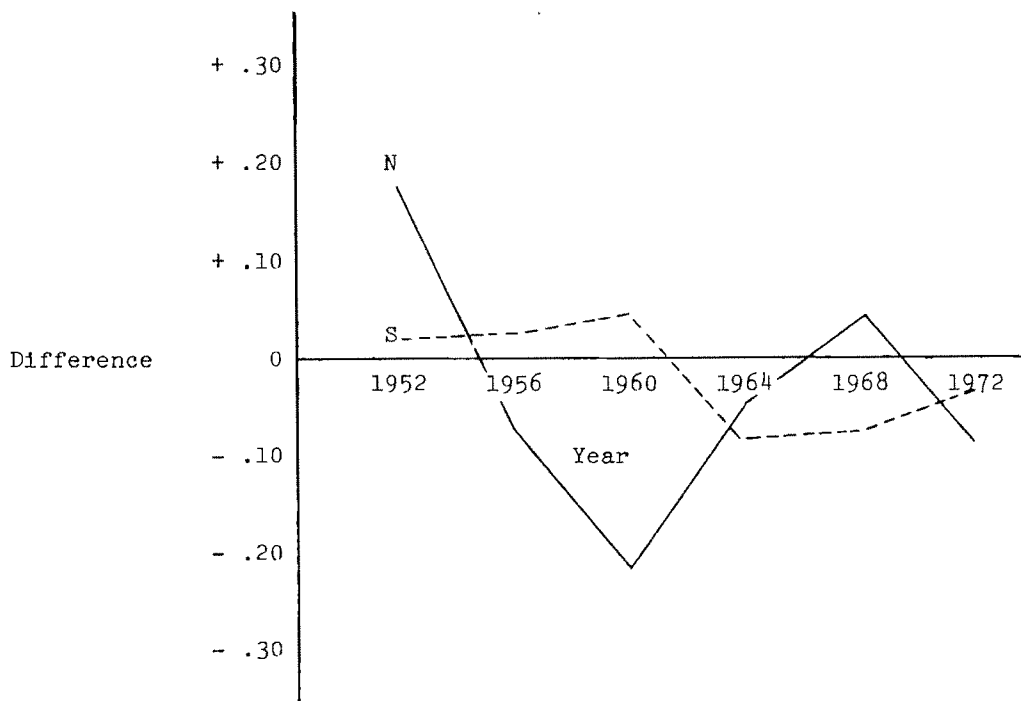


Figure 4. Adjusted Differences in Proportion Participating Plotted by Sign and Magnitude, by Region and by Year

tent with the general working hypothesis of the political climate theory, is the fact that adjusted registration, voting and participation differences between whites and blacks did indeed change over time differently for each region. Furthermore, as expected, both adjusted registration and voting coefficient changes over time were larger in the South than in the North. Finally, as predicted by both hypotheses five and six, 1952 and 1956 were characterized by isolation (11 of 12 coefficients positive) and 1964 by ethnic community (4 of 6 coefficients negative).

Most of what discrepancy exists between the remaining predictions and data boils down to two pairs of unexpected findings which are mostly resolvable in the context of a modification of the political climate theory. First, in the registration and voting figures, the 1968 and 1972 coefficients for the North are very slightly positive instead of the anticipated slightly negative. In the case of 1972 this is no problem since both the expected and actual coefficients are quite close to zero. The 1968 results, however, suggest that the ethnic community evident in 1964 was short-lived in the North. Because the brunt of the attacks on political discrimination was in the South, this is not a strong criticism of the political climate theory. In the South, the registration and voting curves are both negatively linear instead of curvilinear as predicted and the transition period later than expected: 1964-1968 instead of 1960. A plausible explanation for the late transition period consistent with the political climate assumption is that the effect of political discrimination on black political activity is particularistic. That is, while ethnic community might have been manifest in other political activities due to the civil rights movement of the early 1960s (and the results on participation comparisons support this assumption), only when barriers specific to registration and voting were removed (via the Civil Rights and Voting Rights Acts of the middle and late 1960s) could one expect to see ethnic community in effect at the polls. Therefore, the only problem with the political climate argument here is the lack of specificity in indi-

cators of political climate. Civil rights demonstrations clearly did appear to reduce the political isolation of Southern blacks, as the steady reduction in positive coefficients from 1952 to 1968 indicates; but only after legislation aimed directly at removing obstacles to black registration and voting was passed and began to be enforced did blacks vote in greater proportions than whites at the polls.⁵

A second set of unexpected findings relates to the political participation comparisons. In the North, there is much more fluctuation than expected, but, at the same time, only two of the four coefficients are statistically significant. Because the original hypothesis five predictions anticipated coefficients close to zero for several of the years, this may not seem surprising. On the other hand, the strong negative coefficient in 1960 is not explained away by such reasoning. An alternative explanation is to look at the registration, voting and participation results for the North together and note that, out of 18 coefficients, 13 are statistically insignificant, suggesting the indirect effects argument is more appropriate to the political behavior of Northern blacks than is the political climate theory during this time period (although the trends noted for Northern registration and voting comparisons are somewhat consistent with the political climate argument). Quite simply, it is possible the degree of political discrimination assumed to exist in the North during the period 1952-1972 (mild, as predicted in hypothesis six) was low enough to be basically irrelevant.

In the South, the sign of the participation coefficients conform to the political climate predictions, but their magnitudes (save in 1964) are modest. This is understandable in light of the different types of activities which participation represents. Registration to vote and voting, on the one hand, are highly visible, directly threatening forms of political participation, while donating money, talking politics with family or friends, wearing campaign buttons,

⁵ Such an interpretation is consistent with Walton's (1972) theory of black political development as being composed of a series of stages.

going to meetings and working for candidates, on the other hand, are relatively hidden and/or quite nonthreatening forms of participation.

In fact, within nonvoting political activities,

the greatest difference between Southern whites and Negroes in campaign participation is found in the most visible form of activity (attending meetings and rallies), and the smallest difference is found in the least visible form (giving money and buying tickets). (Matthews and Prothro, 1966:49)

The visible forms of participation are those most likely to threaten and, therefore, to conform to the predictions of the political climate theory in which perceived or real political threats of the minority and resultant political discrimination by the majority are fundamental. In fact, the repression of black voting is seen as encouraging less visible forms of political participation (see Matthews and Prothro, 1966:51), suggesting that there is nothing inherently apolitical or apathetic about blacks as a minority group. Therefore, less visible forms of political participation are likely to be less hindered during bad times (isolation) and less helped during good times (ethnic community) than more visible forms of activity like registration and voting.

CONCLUSIONS

Contradictory arguments supported by apparently discrepant findings on the difference between white and black political activity levels from recent studies have been subjected to a multivariate trend analysis based on national presidential election year survey data. The results of registration, voting and other political participation comparisons strongly support a modified political climate theory which is more precise with respect to time, place and type of political activity than the original political climate theory discussed at the outset of this paper. Specifically, as seen particularly in the South, two conclusions are in order. First, a strongly intolerant political climate *does* suppress black registration and voting—two visible and quite threatening modes of political participation—but alters only slightly less

visible and seemingly harmless methods of participation like donating money, talking politics with family, etc. Second, a supportive political climate *does* encourage highly visible (and potentially quite rewarding) forms of political activity like registration and voting but not so much other forms of political activity.

As for the future, the ethnic community seen to be evolving in the South during the late 1960s and early 1970s may be self-generating once firmly established. But it may not, and the new South may go the way of the old North where indirect effects appear to explain what few differences in political activity exist between blacks and whites. Historical data from the colonial period and the post-Reconstruction South also suggest this possible demise or coming hibernation of ethnic community. Contemporary evidence on black office holders and the wooing of black constituencies, on the other hand, is more encouraging. Whatever the future holds, it is likely the answer lies within the dynamic framework of a modified political climate theory where the interrelationship among indirect effects, isolation and ethnic community assumptions can be understood.

REFERENCES

- Agger, Robert E., Daniel Goldrich and Bert E. Swanson
1964 *The Rulers and the Ruled: Political Power and Impotence in American Communities*. New York: Wiley.
- Antunes, George and Charles M. Gaitz
1975 "Ethnicity and participation: a study of Mexican-Americans, blacks and whites." *American Journal of Sociology* 80:1192-211.
- Campbell, David and Joe R. Feagin
1975 "Black politics in the South: a descriptive analysis." *Journal of Politics* 37:129-62.
- Clemente, Frank and William J. Sauer
1975 "Voluntary associations and minority status: comment and extension." *American Sociological Review* 40:115-7.
- Daniel, Johnnie
1969 "Negro political behavior and community political and socioeconomic structural factors." *Social Forces* 47:274-80.
- Danigelis, Nicholas L.
1977 "A theory of black political participation in the United States." *Social Forces* 56: 31-47.

- Downes, Bryan T. and Stephen W. Burks
1969 "The historical development of the black protest movement." Pp. 322-44 in Norval D. Glenn and Charles M. Bonjean (eds.), *Blacks in the United States*. San Francisco: Chandler.
- Flanigan, William H.
1972 *Political Behavior of the American Electorate*. Boston: Allyn and Bacon.
- Glenn, Norval D.
1970 "Problems of comparability in trend studies with opinion poll data." *Public Opinion Quarterly* 34:82-91.
- Gosnell, Harold F.
1935 *Negro Politicians: The Rise of Negro Politics in Chicago*. Chicago: University of Chicago Press.
- Killian, Lewis
1968 *The Impossible Revolution? Black Power and the American Dream*. New York: Random House.
- Lane, Robert E.
1959 *Political Life: Why and How People Get Involved in Politics*. New York: Free Press.
- Lipset, Seymour M.
1960 *Political Man: The Social Bases of Politics*. Garden City: Doubleday.
- Lomax, Louis
1971 *The Negro Revolt*. New York: Harper and Row.
- Matthews, Donald R. and James W. Prothro
1966 *Negroes and the New Southern Politics*. New York: Harcourt, Brace.
- Milbrath, Lester W. and M. L. Goel
1977 *Political Participation*. 2nd ed. Chicago: Rand McNally.
- Olsen, Marvin E.
1970 "Social and political participation of blacks." *American Sociological Review* 35: 682-97.
- Orum, Anthony M.
1966 "A reappraisal of the social and political participation of Negroes." *American Journal of Sociology* 72:32-46.
- Pinkney, Alphonso
1975 *Black Americans*. 2nd ed. Englewood Cliffs: Prentice-Hall.
- Simpson, George E. and J. Milton Yinger
1972 *Racial and Cultural Minorities*. New York: Harper and Row.
- Walton, Haynes, Jr.
1972 *Black Politics*. Philadelphia: Lippincott.
- Williams, J. Allen, Jr., Nicholas Babchuk and David R. Johnson
1973 "Voluntary associations and minority status: a comparative analysis of Anglo, black, and Mexican Americans." *American Sociological Review* 38:637-46.
- 1975 "Compensatory and ethnic community theories: are they generalizable?" *American Sociological Review* 40:117-21.

COMMENTS

BLACK SUBORDINATION AND WHITE ECONOMIC WELL-BEING

(COMMENT ON SZYMANSKI,
ASR JUNE, 1976)*

This comment concerns two articles, only one of which appeared in ASR. But since the article by Szymanski (1976) is simply a less sophisticated quasi-replication of an article by Reich (1971), both must be considered.

Both articles deal with the question of white gain from the subordination of blacks; both reject previous findings suggesting such gain (e.g., Glenn, 1966; Thurow, 1969); and both, arguing from a Marxian perspective, present evidence purporting to document white loss from black subordination. Both articles also suffer from severe conceptual and methodological problems that invalidate their conclusions. These problems are first noted, and then both data sets are reanalyzed. Conclusions opposite to theirs are drawn.

Methodological Problems

Reich's findings are invalidated by the fact that he uses family income in his measures. Since family income results from a plethora of combinations of varying numbers of persons working some full and some part time at different levels, it is a poor (and notoriously misleading) measure of economic discrimination. The black/white family income gap is not a valid indicator of discrimination (or even of earnings inequality). If it were we could end all discrimination by limiting white families to one adult wage earner and requiring all black families to have at least two. The family should be our primary focus, but first we must sort out labor market factors with individual data, then we can ascertain the effect of those factors on family income and thus on the stratification system in general.

This unfortunate choice of a major independent variable is sufficient to call into question all of Reich's findings. But even if his measure were adequate, other problems intrude. To control for SMSA income level, he introduces median white-family income into his equations along with the black/white family income ratio

measure. A zero-order correlation of .832 between these two SMSA measures could signal a problem of collinearity. This would be serious, since Reich's whole analysis revolves around the interpretation of the ratio measure coefficient. Beyond this, the meaning of the coefficient of B/W, net of W, is at least troublesome. Further, Reich (and Szymanski) ignores massive interaction effects, as shown below.

Szymanski avoids Reich's most obvious flaw, employing earnings of persons rather than income of families. Unfortunately, he discards Reich's already inadequate measure (income share received by top 1%) and substitutes a worse one: white male median income. This measure is not very meaningful. The absolute dollar earnings of those at the 50th percentile may be of some interest, but certainly of no greater interest than the absolute earnings of those at the 25th or 75th percentile. More importantly, the dollar value at any particular point on the income curve is not dependably predictive of the curve itself—that is, we cannot safely infer differences at the 55th percentile from differences at the 50th (see Wohlstetter and Coleman, 1972, for an extensive discussion of this point). Szymanski's focus on this particular percentile as a summary of the well-being of most of the income curve (white workers) is clearly inappropriate, as is his use of the Gini in this regard (see below). Further, the unnecessary relating of white median income to the ratio of black/white median income exacerbates the already troublesome problem of a built-in dependency of the variables in question.

A related, but far more serious, problem is Szymanski's (and Reich's) measure of racism. Inherent problems in the use of median ratios for this purpose have been well documented (see Villemez and Rowe, 1975; Lieberman, 1976; Villemez, 1977a; 1977b). The problems all have a central focus: the ratio of two single points on two income curves cannot describe adequately or accurately the relationship of the two distributions. Wohlstetter and Coleman's (1972) demonstration of numerous distinct linear trends across black/white income-ratio at quantiles curves indicates that comparisons at different points would generate quite different findings. It is important to note that the specific shape of ratio at quantiles curves varies from state to state (and from SMSA to SMSA), which makes a median ratio comparison across those units even more questionable.

*Address all communications to: Wayne J. Villemez; Department of Sociology; University of Illinois at Chicago Circle; Chicago, IL 60680.

Given those problems, Szymanski's purported demonstration of an inverse relationship between racism and the economic well-being of white workers boils down to the discovery of an association between, on the one hand, the income of whites at the 50th percentile, and on the other hand, that same income compared to that of blacks at the 50th percentile. And even this unconvincing association turns out to be a statistical artifact. Szymanski uses only partial correlational analysis to demonstrate his point, and the technique is inadequate to the model. In his implicit model Szymanski assumes a great many B_{ij} equal to zero, but makes none of the requisite empirical predictions that follow therefrom (though, in fact, many of the unreported second and higher order partials suggest the rejection of his hypothesis). Phenomena as complex as income distribution are difficult to illuminate with correlational analyses, and simply cannot be examined with only first order partials. Szymanski's test, far from conservative, is not a test at all.

Dealing solely with correlational measures can also lead one into logical conundrums. For example, Szymanski finds that controlling for percent union in a state reduces to zero the correlation between black/white median income ratio and white median income. He concludes: "... [T]he relationship between these two measures is more or less caused by the effect of racial discrimination on working class solidarity and the effect of working class solidarity on white male median earnings" (Szymanski, 1976:411). Leaving aside what "more or less caused" may mean, we have a supposed demonstration of a causal chain (discrimination \rightarrow unionization \rightarrow white earnings) that is actually an *a priori* assumption—with cross-sectional data, it can be nothing else. Yet little supporting argument is offered for this assumption. Ashenfelter (1972), for example, assumes the causal order of the first two variables to be reversed. In point of fact, the relationship among unionization, discrimination, and white income is enormously complex, and it is recognized as such in the radical economic literature. Baron (1975), for example, has traced the historical use of black labor by both white capital and white labor, and has pointed out historical instances in which working-class groups as well as employers and the ruling class have used racism to their advantage. Racism is not always to the interest of the capitalist class—it has been often to the interest of the working class as well. The question of racial solidarity vs. class solidarity is a difficult one and is a question upon which Marx was vague and Engels, wisely equivocal. The

use of a gross percentage-unionized measure as a surrogate for class solidarity is unacceptable and will be omitted herein.

These and other methodological problems are sufficient to make the white-loss findings of the radical economists and their devotees suspect, but other types of problems also intrude.

Conceptual Problems

Reich takes the absolute income of black families at the 50th percentile as a percentage of the absolute income of white families at that percentile, and says he is measuring racism. Szymanski terms his similar median ratio a measure of the intensity of racism. Both are obviously measuring income inequality of some sort, but not necessarily even discrimination, much less racism. Not all inequality results from discrimination, and not all discrimination against blacks derives from racism (for a fuller discussion of this point, see Villemez, 1977b). The measure employed herein is a measure of black/white income inequality superior to those used before. But to call it discrimination, nonetheless, would be an assumption, signaling the need for cautious interpretation. This is not a trivial distinction. To speak of white gain from discrimination is one thing; to speak of white gain from inequality quite another. In the latter case the gain is inextricably linked (in a measurement sense) to the inequality, despite the fact that gain on other whites vs. inequality of blacks and whites is at issue. The question becomes recast as: In areas where whites are better off relative to blacks, are they, *ceteris paribus*, better or worse off relative to whites in areas with less of a differential? Following that, the next logical question is: Why is that? Calling a measure of inequality discrimination answers the second question by fiat.

Misrepresenting some previous literature in the process, Szymanski greatly oversimplifies the problem. His most egregious misreporting is of Glenn's (1966) seminal article. Glenn's paper is concerned as much with nonmonetary white gains as with monetary gains—his cautious conclusions clearly reflect this focus. Szymanski cites only those portions dealing with monetary gain. Both he and Reich focus only on income gains to whites from black/white income disparities. To focus solely on seeking out the effect of the relationship between black and white aggregate income on white income is not an entirely reasonable enterprise. Economic inequality is only roughly and partially captured by a comparison of two income distributions. Income inequality is not a full measure of economic inequality. Put an-

other way, white income gain at the expense of blacks is not necessarily a direct quid pro quo. To expect the relationship between black and white *income* to predict adequately white *income gain* is an oversimplification. If whites do in fact benefit in terms of income from black economic subordination it is indirectly through a complex of historical factors and labor market dynamics. We can, however, validate the continuing effort to understand those dynamics by demonstrating that the two recent analyses notwithstanding, white workers do not lose from discrimination, and, indeed, seem to be better off in those areas where blacks are relatively worse off.

Reanalysis

The data reanalyzed are the same state-level statistics employed by Szymanski (1976), as well as the SMSA data of Reich—Reich's 1960 data were for 48 SMSAs; herein 1970 data for all SMSAs over 250,000 are employed. SMSA data are clearly superior (since SMSAs approximate labor markets, states do not), but both types initially are used to illustrate that methodological and conceptual problems are at issue, not questions of unit of measurement.

The measure of black/white income inequality used is the Index of Net Difference (ND) developed by Lieberman (1976).¹ Aside from ND and the dependent variables (see below), the same measures used by Reich and Szymanski are employed, for the SMSA and state data, respectively. Two separate measures of white economic well-being are employed: the proportion of white workers earning over \$15,000 a year, and the proportion earning under \$5,000 a year. The Gini concentration ratio used by both Reich and Szymanski is not included since it is a notoriously poor measure of the well-being of any segment of the curve (see Bronfenbrenner, 1971; Paglin, 1975; Villemez and Wiswell, 1978), and since exact proportional equality is not of itself an economic benefit. The two measures in tandem should give a somewhat clearer picture of the size distribution of white male income than a single summary measure could. The dependent variables were each regressed on ND_{bw}, percent black, and percent of the labor force in manufacturing. This was done for the non-South and South, using SMSA data and state data. Indication of white gain would be positive

Table 1. Net Effects of Black/White Inequality, Percent Black, and Percent of the Labor Force in Manufacturing on the Proportion of White Males Earning over \$15,000 and under \$5,000, for Non-Southern SMSAs and States

Dependent Variable	ND _{bw}	% Black	% in Manufacturing	R ²
Non-South SMSA				
15+	.04	.33*	-.04	.31
5-	-.08	-.40*	-.28*	.37
State				
15+	.04	.75*	-.04	.35
5-	-.11	-.71*	-.31*	.59
South SMA				
15+	.08	.02	-.12	.12
5-	-.32*	-.08	-.17	.33
State				
15+	.68*	-.66*	-.29*	.66
5-	-.45	.35	.16	.32

* Double its standard error.

partial coefficients for the \$15,000+ variable and negative coefficients for the \$5,000- variable. Such a pattern for ND_{bw} would imply that net of percent black and percent in manufacturing, the higher the black/white income inequality, the greater the percentage of whites earning over \$15,000 and the smaller the percentage earning under \$5,000. The same pattern for percent black would show similar white advantage in areas with higher black concentrations.²

The raw findings are presented in Table 1. As is evident from the signs of the coefficients, whites at all levels seem to gain from the presence of blacks and from black subordination. With only one inconsistency (percent black, South, state data), the greater the black/white inequality, the more whites over \$15,000 and the fewer under \$5,000; also, the greater the proportion of blacks in an area, the more whites over \$15,000 and the fewer under \$5,000. This finding directly contradicts those of Reich and Szymanski.³ However, since few of the ND_{bw} coefficients are statistically sig-

¹ Applied to income, ND provides a clearer measure of the extent of black male income inequality (or more precisely, a measure of the additional probability of higher income that follows from not being black). There are problems with this measure (see Waldman, 1977) but it is superior to most.

² In this comment I retain the focus of Reich and Szymanski on benefits to white workers only. With their data (and thus mine), the question of whether white capitalists (owners, employers) also benefit, cannot be considered. Blacks only are considered here rather than blacks plus Spanish (Szymanski's third-world). Equations containing both groups were run as a safeguard, with no substantive differences resulting.

³ Inserting white male median income into the equations as a control measure of areal economic status is a questionable procedure. For what it is worth, these findings hold even when that variable is added.

Table 2. Unstandardized Net Effects from an Interaction Model, for High-Low Percent Black and High-Low Percent in Manufacturing Areas, for SMSAs, for the Non-South and South.

Dependent Variable (% of white males earning \$15,000+ or \$5,000-)	ND _{bw}				% in Manufacturing		% Black		% Black	% Manu
	High Blk	Low Blk	High Manu	Low Manu	High Blk	Low Blk	High Manu	Low Manu		
Non-South										
15+	.27*	.01			.09	-.07			-.11*	
15+			.01	.14*			.41*	.25*		.02
5-	-.35	-.05			-.45*	-.23*			.14	
5-			.01	-.53*			-.63*	-.19		-.19
South										
15+	-.03	.14			-.15	-.04			.11	
15+			.25	.09			.25	.01		-.12
5-	-.24	-.36*			-.14	-.26			-.10	
5-			-.98	-.30*			-.99	-.09		.42

* Double its standard error.

nificant, all we can say with assurance is that there is no evidence whatsoever of white loss from black economic subordination, and there is clear indication of white economic gain from the presence of blacks.

But we can go further than this. Previous research would suggest considerable interaction among the independent variables, and simple tests for interaction show that there are several such confounding the results. Only one of the several equations is presented as example. Where $X_1 = ND_{bw}$; $X_2 =$ percent black; $X_3 =$ percent in manufacturing; and $Y =$ percent whites earning over \$15,000, we find for non-southern SMSAs:

$$\hat{Y} = .07 + .52X_1 + 2.33X_2 + .39X_3 - 6.82X_1X_2 - 1.46X_1X_3 - 6.95X_2X_3 + 23.40X_1X_2X_3 \quad (R^2 = .41).$$

(*Double its standard error.)

The coefficients of both the three first-order interactions and the second-order interaction are significantly different from zero, and the explained variance is increased by 1/3 (10%). This pattern holds across most equations.

Fortunately, the task of sorting out such complicated conditional relationships has been greatly simplified by Wright (1976). His elaboration of Gujarati (1970), while an elegant technique, is simple in concept. By simply extending normal dummy variable scoring, one estimates a model: $Y = C_0 + C_1X^{(S)} + C_2X^{(N)} + C_3G^{(S)} + e$; where C_0 is the intercept for the base category, group N; C_1 is the slope of X for group S; C_2 is the slope of X for group N; C_3 is the difference between the intercept of group S and the intercept (C_0) of group N; $X^{(S)}$ equals X for cases in group S and 0 otherwise; $X^{(N)}$ equals X for cases in group N and 0 otherwise.⁴

⁴ As Wright (1976) notes, the model is probably

Applying this method in a straightforward manner to the SMSA data, we derive the findings in Table 2.⁵ There we find clear evidence of white gain from the subordination of blacks, especially in the non-South. The gain from income inequality is primarily in high percent black areas (coefficient of $ND_{bw} + .27$ for 15+, $-.35$ for 5-) and low percent in manufacturing areas ($+.14$ for 15+, $-.53$ for 5-). The coefficients of percent in manufacturing show white gain from this primarily at the bottom of the income distribution (\$5,000-), and greater white gain from it in the presence of higher concentrations of blacks ($-.45$ vs. $-.23$). The coefficients of percent black show white gain at the top and bottom of the income distribution from the presence of blacks; greater gain at both ends in high manufacturing areas ($-.63$). All relevant coefficients are statistically significant and directly contradict both the findings and the inferences of Reich and Szymanski. Findings for the South are more equivocal, but generally supportive of a white gain hypothesis. More could be inferred from this table—the relationships are incredibly complex—but these few points suffice for the issue at hand.

In sum, it does not necessarily follow from any theoretical perspective that a white loss should result from discrimination against

most clearly understood as two separate regressions within a single equation: Y regressed on X among group N, and Y regressed on X among group S. That is, $Y_N = a_N + C_2X^{(N)} + e$, and $Y_S = a_S + C_1X^{(S)} + e$; where $a_N = C_0$ and $a_S = C_0 + C_3$ (see Wright, 1976, for an explanation of the benefits of using a single equation, and for elaboration and extension).

⁵ Percent in manufacturing was dichotomized at 35%, percent black at 10%, capturing the upper quartile of each distribution.

blacks. Further, convincing evidence from prior research indicates the exact opposite to be the case. Radical analyses purportedly demonstrating such loss have suffered from weak methodology and oversimplification. The reanalysis of data herein has demonstrated that whites clearly do not lose from the economic subordination of blacks, and in many areas they just as clearly gain therefrom. The earlier findings of Glenn, Thurow, and others seem to be substantially correct.

Wayne J. Villemez
University of Illinois, Chicago Circle

REFERENCES

- Ashenfelter, Orley
1972 "Racial discrimination and trade unionism." *Journal of Political Economy* 80:435-64.
- Baron, Harold M.
1975 "Racial domination in advanced capitalism: a theory of nationalism and divisions in the labor market." Pp. 173-216 in R. C. Edwards, M. Reich, and D. M. Gordon (eds.), *Labor Market Segmentation*. Lexington: Heath.
- Bronfenbrenner, Martin
1971 *Income Distribution Theory*. Chicago: Aldine.
- Glenn, Norval
1966 "White gains from Negro subordination." *Social Problems* 14:159-78.
- Gujarati, Damodar
1970 "Use of dummy variables in testing for equality between sets of coefficients in two linear regressions: a generalization." *American Statistician* 24:18-22.
- Liebertson, Stanley
1976 "Rank-sum comparisons between groups." Pp. 276-91 in David R. Heise (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.
- Paglin, Morton
1975 "The measurement and trend of inequality: a basic revision." *American Economic Review* 65:598-609.
- Reich, Michael
1971 "The economics of racism." Pp. 80-92 in Martin Carnoy (ed.), *Schooling in a Corporate Society*. New York: McKay.
- Szymanski, Albert
1976 "Racial discrimination and white gain." *American Sociological Review* 41:403-14.
- Thurow, Lester
1969 *Poverty and Discrimination*. Washington, D.C.: Brookings Institution.
- Villemez, Wayne J.
1977a "The functional substitutability of blacks and females in the labor market: a closer look." *Sociological Quarterly* 18:548-63.
1977b "Male economic gain from female subordination: a caveat and reanalysis." *Social Forces* 56:626-36.
- Villemez, Wayne J. and Alan R. Rowe
1975 "Black economic gains in the sixties: a methodological critique and reassessment." *Social Forces* 54:181-93.
- Villemez, Wayne J. and Candace H. Wiswell
1978 "The impact of diminishing discrimination on the internal size distribution of black income: 1954-1974." *Social Forces* 56:1019-34.
- Waldman, Loren K.
1977 "Types and measures of inequality." *Social Science Quarterly* 58:229-41.
- Wohlstetter, Albert and Sinclair Coleman
1972 "Race differences in income." Pp. 3-81 in A. H. Pascal (ed.), *Racial Discrimination in Economic Life*. Lexington: Heath.
- Wright, Gerald C., Jr.
1976 "Linear models for evaluating conditional relationships." *American Journal of Political Science* 20:349-73.

WHITE WORKERS' LOSS FROM RACIAL DISCRIMINATION: REPLY TO VILLEMEZ*

Central to Villemez's argument is his claim that the Index of Net Difference (ND) is superior to the ratio of median incomes as a measure of inequality. While this is true for some purposes requiring an indicator of the *whole income distribution*, it is not true in studies such as mine which focus on working-class income.

In making his argument Villemez misrepresents the sources he relies on. Liebertson (1976:284-5) argues,

... the index of net difference should be viewed not as an alternative to the median but as an alternative to the use of medians in studies that are more centrally concerned with a comparison between entire distributions. ... If there is intrinsic interest in ... the location of the 50th percentile ... then the median should be used to compare two or more populations.

Liebertson argues that there usually is not a radical difference between the two measures, e.g., he computed that for the years between 1947 and 1968 the correlation between them was +.89 (1976:284).

Villemez sights the careful analysis of the pitfalls of comparing income distributions by Wohlstetter and Coleman (1972). In fact, these

*Address all communications to: Albert Szymanski; Department of Sociology; University of Oregon; Eugene, OR 97403.

I wish to thank Larry Carter, Harry Humphries and Michael Reich for their helpful comments on an earlier draft of this article. I would especially like to thank Gerald Duguay for his help in making this response much stronger than it otherwise would have been.

authors showed that in 1967 the ratio of non-white to white median income was a valid predictor of the ratio of incomes from the thirtieth to the ninetieth quantile—the ratio over this entire range was approximately .65 (1972: Fig. 1–4). In 1969 the median income in the urban United States for men was \$6,860, while that of male craftsmen was \$7,693 and factory operatives \$6,536.¹ The median ratio, at least of personal income in the late 1960s, thus is seen to be an excellent measure of the economic position of working-class whites compared with blacks.

Villemez also cites Wohlstetter and Coleman's argument that in some comparisons separated by only a few years, when there are relatively small changes in relative income, a comparison of medians can indicate no change while a comparison of quantiles less than the median indicates either a slight relative decline (e.g., 1949–1959) or improvement (e.g., 1945–1952). A comparison of quantiles above the median indicates the opposite. This is because of contrary factors operating on low and high income nonwhites. While certainly an interesting (and atypical) phenomenon, and a warning against making sweeping statements about income changes on the basis of a change in a few percentage points at any single measurement point, these findings are not relevant to Villemez's conclusion. First, the difference between the quantile ratio curves for 1949–1959 averaged only .02 and that for 1945–1952 only .05. In both cases the differences were as much above as below. When one considers the level of sampling error (the data are based on a sample of only 8,000 households), these slight differences are not large enough to make a valid claim for anything other than there being no significant change between the two years (1972: Figs. 1–17, 1–18). Second, in comparisons of either the states or the SMSAs, since they include far more cases than the comparisons between two years, unique effects such as those noted by Wohlstetter and Coleman are much less likely to occur since any one unit's peculiar distribution is very likely to be cancelled out by those of others.

While the ND for some purposes is superior as a measure of *overall* inequality, whether or not this is so in any specific case depends on the formulation of the problem to which it is applied. The ND has two component parts: (1) the probability that a white selected at random will be earning more than a black similarly selected, and (2) the probability that a black selected at random will be earning more than a

white similarly selected. Lieberman's ND is (1) minus (2). This expression thus varies from +1.0 where all whites earn more than all blacks to –1.0 where all blacks earn more than all whites. In other words, extremes of racial inequality can either have (+) or (–) values while racial equality is manifested in scores near zero. This is fine for studies interested only in the correlates of how much whites are ahead of blacks, but is most confusing in studies such as mine which are interested in the degree of inequality between two populations without regard for the direction of that inequality. My thesis can be stated as follows: the greater the inequality between two ethnic groups within the working class, the more disunity within the working class and, hence, the worse off *both* ethnic groups. The ND is incapable of accurately measuring such inequality. This is because it subtracts the relative advantage of blacks from that of whites *before* entering the components of ND into the regression equations, rather than entering both terms separately as must be done to detect the net effects of *both* components.

Specifically, where we expect antiblack racism to predominate we should focus on the first component of ND, the probability that a white will be earning more than a black, and we should examine independently the net effects of this factor. The Villemez model using ND predicts that a high positive ND should be associated with a high percentage of whites over \$15,000, while a high negative ND should be associated with a low percentage of whites over \$15,000. My model, on the other hand, predicts that either a high positive or a high negative ND would correspond to a low percentage of whites over either the median or over a set figure such as \$15,000 (if we are concerned about the entire white distribution). Use of only the probability that a white selected at random would be earning more than a black similarly selected in place of the ND gives us a consistently different prediction. Villemez here would predict that the higher this figure the higher the percentage of whites earning over \$15,000, while my model predicts the higher this figure the lower the median income (or percentage over \$15,000) among whites (if my model is applied to the white distribution *as a whole* rather than just to the white working class).

Villemez's regressions were rerun entering both terms of ND *separately* into the equations (using the 50 states as the units and running the regressions both with and without the dummy variable of region). When this is done his results are in all cases *reversed*. When the percentage of whites earning over \$15,000 is re-

¹ All data on occupational earnings are from the U.S. Department of Commerce, 1971:229.

gressed on the probability of a white earning more than a black, the net effects are *negative*. They are positive when the percentage of whites earning under \$5,000 is regressed on the probability of a white earning more than a black. (In all four cases the results are significant at double their standard errors.) A more refined use of Lieberman's measure thus shows the opposite of what Villemez reports on the basis of ND. In fact, not only the working-class whites, but whites *as a whole* lose economically from racial discrimination (with the exception of the very highest income whites not adequately reflected in general income distributions).

In summary, although in some contexts the ND is a slightly better measure of overall relative inequality than median ratios for skewed distributions, and the probability of a white selected at random earning more than a black similarly selected is generally a better measure than the ND for studies of the divisive effect of racism on entire distributions, the use of median ratios is almost always a fairly good measure of overall inequality in skewed distributions. As a measure of working-class people's position, it is superior to both of the former measures.

Villemez summarily dismisses the Gini index of inequality on the grounds that (1) the Gini is a distribution-free statistic and (non sequitor) therefore measures nothing of interest to those analyzing the effects of discrimination on white income; and (2) income equality in itself is not a benefit to whites. I certainly can not agree more with the latter point; income equality is only a benefit to those with lower incomes (the people focused on in my argument). Villemez's former point overlooks the fact that he makes precisely the opposite claims in dismissing the Gini that he makes in *defending* the ND. The alleged superiority of the ND lies in its being distribution free, the same characteristic that makes the Gini unacceptable to Villemez. In fact, as Lieberman (1976:281) points out, the Gini and the ND are analogous measures. What the Gini is to a single population distribution, the ND is to the comparison between population distributions.

Both the Gini and the ND can be valid measures of the degree of *aggregate* inequality, while neither is a precise predictor of any particular quantile. A lower ND score (or a lower probability that a white selected at random will be earning more than a black similarly selected) indicates that as a whole blacks are better off in comparison with whites (without predicting how much better off is any particular quantile). A lower Gini indicates how much better off lower income people are, in relation

to higher income people (without predicting how much better off is any particular quantile in comparison with any other given quantile). It would seem that all the strengths and weaknesses of the ND are also the strengths and weaknesses of the Gini.

My results using the Gini as a measure of the degree to which lower income whites benefit in relation to higher income whites stand—the *less* the discrimination against blacks, the *less* the inequality among whites (i.e., poorer whites benefit from lesser discrimination against blacks). The greater the proportion of an area that is Third World, the more inequality among whites (i.e., poorer whites lose from the presence of large numbers of Third World people). The same results also are found when the white Gini is regressed (using the 50 states) on either the ND or the probability that a white selected at random is earning more than a black similarly selected. In both cases the results are significant at double their standard errors.

Villemez argues that the correct units for analyses such as ours are the SMSAs, because they, unlike the states, are common labor markets. While SMSAs are probably slightly better units of comparison because they are more common labor markets than states, their rejection as valid comparative units is not justified. States, too, can be considered common labor markets. For one thing, people move around relatively freely and can easily commute over a significant section of the densely populated areas of most states. For another thing, it is not necessary for every individual to compete for every job (only that there be sufficient mobility to exert pressure on all jobs). Second, the importance of economically coherent units for comparative analysis is exaggerated. The assumption in comparative analysis is that the variables are interrelated by common underlying forces which produce the variation. Thus, industries, occupations, counties, townships, regions or even arbitrary grids of so many square miles are (and often have been used as) valid units for comparative purposes so long as there is a significant variation among the variables to be compared among the units. The fact that units of comparison do not make much difference in the results obtained (so long as the number of units is large enough to cancel out unique factors) is born out by Villemez's own results. He found a net effect of +.04 between the ND and the percentage of whites earning over \$15,000 (non-South) with *both* the states and the SMSAs as his units (see Villemez's Table 1).

Villemez substitutes for my use of the median as an indicator of white gain, the percentage of white males earning over \$15,000 and

those earning less than \$5,000. Again, in his haste to dismiss my analysis, he makes contradictory arguments. Medians are dismissed because they allegedly merely measure the condition of the 50th quantile, but the percentage over \$15,000 (or under \$5,000) is supposed to tell us more than merely the condition of whites at the \$15,000 (or \$5,000) point. Just as there are an infinite number of distributions with the same median, there are also an infinite number of distributions with the same percentage over \$15,000. We can infer no more from an absolute measure such as over \$15,000 about an entire income distribution than we can from a relative measure such as the median. Whether over \$15,000 (or under \$5,000) or the median is a better measure depends solely on whether or not there is an inherent interest in the middle income ranges, in which case the median is the superior measure, or in the higher and lower extremes, in which case Villemez's are better. Since in 1969 only 10.8% of urban white males had an income greater than \$15,000 and 35.4% an income less than \$5,000 while the typical income of factory workers was close to the overall median, the median is clearly the superior measure of the potential gain or loss of *working-class* people.

Taking all of the interaction effects into account, as in my original analysis (which reported the fifth order partials whenever any of the first or higher order partials differed significantly from the zero-order correlations) shows that the positive relation between black/white median income and white median income was reduced but still slightly positive when all five control variables were controlled for (+.06). The relationship between percentage Third World and white male median earnings became even more negative (-.25) when all five factors were controlled for simultaneously than in the case of the zero-order relation. In both cases, with all of the interaction effects, including region, controlled for, whites, if anything, are seen to lose from discrimination. In both the cases of the effect of percentage Third World and black/white median income ratios, controlling for all factors simultaneously makes no difference in the size of the Gini coefficient. Likewise, in both cases, with all the interaction effects controlled for, whites have a greater degree of inequality where the impact of discrimination against Third World people is the greatest. In other words, poorer whites *lose* from the presence of Third World people.

Both Villemez and I control for the effect of region (South, non-South). But because of the higher level of racism in the South together with the lower income levels for both races in this region, this presents serious methodolog-

ical problems. If percentage in manufacturing and income per capita are controlled for, controlling for region in addition primarily amounts to controlling for the intensity and potential impact of discrimination—our independent variables. Exclusively looking at the South and non-South separately, as Villemez does, thus amounts to examining the effect of the independent variables under conditions where these same independent variables have in good measure had their effects removed. Quite naturally, one would expect to find at best a weak relationship between the measures of racism and its effects under such conditions. Villemez's conclusions must then become suspect.

Villemez's regressions were rerun for the country as a whole (the 50 states) with the region variable left out. When this is done there is a reversal of his results (for the two regions taken separately). Here it was found that the net effect of ND on the percentage over \$15,000 was *negative* while the net effect of ND on percentage under \$5,000 was *positive* (both significant at double their standard errors). When ND is correlated with region it can be seen that there is an extremely high positive relation with virtually all the Southern states having higher ND scores than the non-Southern states. To control for region as Villemez does results in net effects from which most of the effect of ND is removed.

My original finding that there is a +.16 correlation between white male median income and the ratio of black to white median income when per capita income is controlled for (using the 50 states), and a +.27 relation when percent in manufacturing is controlled for are more valid findings than the +.09 relations found when controlling for region (see Szymanski, 1976a: Table 1).

Villemez reports his findings controlling only for percent black (and not percent Third World) and does a separate analysis of only those SMSAs with the highest percentage of blacks (and not those that have the highest percent Third World). This is incorrect. More Latins are concentrated in dirty work occupations than blacks (outside the South); black median income is higher (in 1969 it was \$5,587 for black males; \$5,156 for Puerto Rican men; and \$4,839 for Mexican men). Further, there has been a relative slippage in the position of Latins relative to blacks since 1950 (Szymanski, 1975). Latins have come to play a central role in the lower reaches of the U.S. economy and to suffer the same types of discriminations and humiliations as blacks and thus must be treated like blacks in studies of the effect of discrimination on whites. Using

the percentage Third World instead of the percentage black is especially important since there is a strong *negative* correlation between percent black and percent Latin in the SMSAs outside the South. This negative relation was so strong that it resulted in a slight negative correlation between the percent Third World and the percent black ($-.06$) in 1969.

Villemez argues that the effect of ND in producing a higher percentage of whites earning more than \$15,000 (and more than \$5,000) indicates that whites gain economically because of discrimination against blacks. In fact, however, the relationship he reports is a spurious one due mostly to the correlation of both percentage of whites over any set figure and the ND or the ratio of black to white median income with the per capita wealth of a SMSA. When the SMSAs of over 250,000 outside the South are analysed with per capita income controlled, the entire relationship between whites earning over \$10,000 and the ratio of median incomes disappears, while the size of the positive correlation between white median income and the median ratio is decreased by half (see Table 1). In my reanalysis of Villemez I used percentage of whites over \$10,000, rather than over \$15,000 so as to take into account the highest paid segments of the working class. (In 1969 29.7% of white urban males had an income over \$10,000).

The correlation in 1969 between percentage Third World and per capita income in non-Southern SMSAs was $+.33$. Thus it is not surprising to find, as Villemez does, that the higher the percentage of blacks in an SMSA the more whites who earn over \$15,000 (and over

\$5,000). The percentage of *blacks* earning over \$10,000 also relates positively to the percent Third World ($+.29$). If Villemez's argument that whites gain from discrimination is valid it must also follow that *blacks* too must gain from discrimination against themselves. That these correlations reflect the $+.82$ correlation between PCI and percent of whites over \$10,000, and the $+.59$ correlation between PCI and percentage of blacks over \$10,000 and not the fact that whites and blacks gain from discrimination against blacks (or Third World people) is borne out when income is controlled for (see Table 1).

Table 1 reports the results of a reanalysis of Villemez's data using the non-Southern SMSAs. Here we see that there is in fact a positive correlation between percent Third World and whites earning over \$10,000 even when PCI is controlled for (although the relationship is weakened). While there is a weak positive relation between percent Third World and white median income, this relationship entirely disappears when per capita income is controlled for. Thus it appears that the middle class and a few highly paid skilled workers (those over the seventieth quantile) benefit from having more Third World people around while most white workers do not. When PCI is controlled, there is no relationship between whites earning over \$10,000 and the ratio of black/white median income, but there is a slight ($+.11$) relationship between this ratio and white median income. Again, this relationship indicates that the bulk of Northern white working people lose somewhat from economic discrimination against blacks (and lose more

Table 1. Discrimination and White Workers' Loss (Product-Moment Correlation Coefficients)

	Northern SMSA's of Greater than 250,000 Population in 1970*			
	Zero Order		Controlling for Per Capita Income	
	Percentage of White Males Earning Over \$10,000	White Male Median Income	Percentage of White Males Earning Over \$10,000	White Male Median Income
Percent Third World**	$+.45$	$+.22$	$+.34$	$-.02$
Percent Black	$-.05$	$-.02$	$-.03$	$+.02$
Black/White Median Income	$+.16$	$+.21$	$+.01$	$+.11$
The Half of SMSAs with the Most Third World People ($>10.5\%$)				
Percent Third World	$+.06$	$-.13$	$+.07$	$-.19$
Black/White Median Income	$+.16$	$+.23$	$-.02$	$+.11$
The Half of SMSAs with the Most Blacks ($>6.4\%$)				
Percent Black	$-.15$	$-.11$	$-.06$	$-.03$
Black/White Median Income	$+.03$	$+.09$	$+.12$	$+.16$

* All SMSAs excluding those in the states of the old Confederacy, Kentucky (a border state), New Mexico (the equivalent of the South for Chicanos) and Hawaii (a predominantly Asian state).

** The total of black and persons with either Spanish surnames or who speak Spanish.

Source: U.S. Bureau of the Census; 1970.

than middle-class people). When PCI is controlled for this relationship is stronger.

When only half of the non-Southern SMSAs with the highest percentage Third World are examined, the relationships reported above between the median income ratios and both measures of white gain stand virtually unchanged while the observed relationship disappears in the case of percentage of whites over \$10,000, and reverses in the case of white median income. This indicates that the middle class does not gain from having more Third World people around and that working-class whites clearly lose. Even though Villemez's use of percent black in place of my percent Third World is conceptually inadequate, in an attempt to replicate his findings I examined the effect of the independent variables in the half of the SMSAs with the highest percentage of blacks. In no case could I find results such as those reported by Villemez. Although the coefficients are small, in all cases the higher the percentage of blacks, the lower the percentage of whites earning over \$10,000 and the lower the white median income (see Table 1). Whether or not we look at all the SMSAs or just those with the highest concentration of blacks, if anything, the higher the percentage of blacks, the lower the white median income and the fewer whites earning over \$10,000. Likewise, whether or not we look at all SMSAs or only those with the most blacks, and with or without controlling for PCI, there is a universal positive association between the ratio of black to white median income (albeit small) and both measures of white gain.

In summary, then, the contrary findings reported by Villemez are a product of his serious conceptual problems, inappropriate statistical measures and selective reporting of data. The mass of evidence, whether obtained from the analysis of states or SMSAs, whether from the whole country or the non-South, whether measured by the white Gini index, the white median income, the percentage of whites over \$15,000, over \$10,000, or under \$5,000, whether measured by the ratio of black to white median income, the Net Difference, the probability of a white selected at random earning more than a black similarly selected, the percent black or the percent Third World, all point in the same direction.² White working

people do *not* gain economically from discrimination against Third World people. In fact, they lose.³

Racism historically has played a central role in American society. It is probably the single greatest reason why the U.S. alone of all the advanced capitalist countries has never developed a massive Communist, Socialist or Labor party rooted in its working people. Since the early days of industrial capitalism in the U.S., one ethnic group has been played off against the other: first the native-born against the Irish, then the English-speaking against the new immigrants from Southern and Eastern Europe, then, especially since World War I, whites against blacks, and most recently, Anglophones against Latins. Racial and ethnic group hostilities and antagonisms rooted in the appearance that one group is getting a better deal than others (slightly better jobs, slightly better pay, special social privileges or fewer burdens of humiliation) encourage the slightly better off groups to act to "protect" their "white skin privilege" against the apparent "threat" from lower status groups. Meanwhile, the poorer and ethnically distinct sections of the working class have borne resentments against the slightly more privileged white, the native-born or the English-speaking, and have tended to blame them for their plight. The net result of this mutual antagonism is an especially low level of working-class consciousness, weak unions, little class solidarity or effective class action. That the divisive effect of racism operates to keep working-class whites poor is no mere speculation of mine. C. Vann Woodward (1974) in his important work *The Strange Career of Jim Crow* documents that the "Jim Crow" system in the U.S. South was developed in the 1890s as a response to the potential for unity between the poor whites and the poor blacks in this region manifested in the growth of the People's party which attempted, almost successfully, to challenge the political and economic domination of the wealthy whites in the South. The Jim Crow system, race baiting, and the myth of white skin

assessed with proving my conclusions wrong, presents a pot pourri of ad hominem arguments, inconsistencies and non sequiturs, misrepresentations of central points in my argument as well as of other people's findings, methodological oversights and errors, faulty conceptualizations and above all selective and misleading reporting of data to attempt to demonstrate conclusions opposite mine.

³ For a further development of the argument that white working people lose because of racial discrimination, the interested reader should refer to two forthcoming publications by Michael Reich (1978; 1979).

² The critique of Villemez developed in this article applies with some minor modifications to his criticisms of my other two articles on the interrelationships among white, male, black and female income. See my articles (Szymanski, 1976b and 1977) and Villemez's critiques (Villemez 1977a and 1977b). In all three of his responses Villemez, apparently ob-

privilege have been *the* central pillars of conservative upper class domination in the South since the 1890s and more than anything else explain the general conservatism, backwardness, lack of unions and relative poverty of white working people in the region. The effect of racism on working-class whites is not limited to the South.

Things are not always what they appear to be, nor are they what powerful interests in the U.S. would like the American working class (black, Latin and white) to believe. White workers are neither the beneficiary nor the cause of racism. We must look elsewhere if we want to understand the causes of racism and thus how to eliminate it. The consequence of understanding that the divisive effects of racism result in whites being poorer than they would otherwise be is that a movement to overcome racism can be based on a coalition of working-class whites, blacks and Latins against the real beneficiaries.

Albert Szymanski
University of Oregon

REFERENCES

- Liebertson, Stanley
1976 "Rank-sum comparisons between groups." Pp. 276-92 in David R. Heise (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.
- Reich, Michael
1978 "Who benefits from racism?" *Journal of Human Resources*. In press.
1979 *Racial Discrimination and Capitalism*. Princeton: Princeton University Press. In press.
- Szymanski, Albert
1975 "Trends in economic discrimination against blacks in the U.S. working class." *Review of Radical Political Economics* 7:1-21.
1976a "Racial discrimination and white gain." *American Sociological Review* 41:403-13.
1976b "Racism and sexism as functional substitutes in the labor market." *Sociological Quarterly* 17:65-73.
1977 "The effect of earnings discrimination against women on the economic position of men." *Social Forces* 56:611-25.
- U.S. Bureau of the Census
1970 *Census of Population*. Washington, D.C.: U.S. Government Printing Office.
- U.S. Department of Commerce
1971 *Statistical Abstract of the U.S.* Washington, D.C.: U.S. Government Printing Office.
- Vann Woodward, C.
1974 *The Strange Career of Jim Crow*. New York: Oxford.
- Villemez, Wayne J.
1977a "The functional substitutability of blacks and females in the labor market: a closer look." *Sociological Quarterly* 18:548-63.
1977b "Male economic gain from female subordination: a caveat and reanalysis." *Social Forces* 56:626-36.
- Wohlstetter, Albert and Sinclair Coleman
1972 "Race differences in income." Pp. 3-82 in A. H. Pascal (ed.), *Racial Discrimination in Economic Life*. Lexington: Heath.

ITEMS (Continued)

■ **CHARLES R. TITTLE** (The Myth of Social Class and Criminality) is Professor of Sociology at Florida Atlantic University. He is doing research on the effect of sanctions on human behavior. Currently he is preparing a monograph entitled *Deviance and Consequences*, which is based on survey data from three states. **WAYNE J. VILLEMEZ** is Associate Professor of Sociology at the University of Illinois, Chicago Circle. His research includes an examination of black/white differences in the residential return on individual resources and an examination of the impact of contextual factors on the relationship between objective and subjective class. **DOUGLAS A. SMITH** is a graduate student in Sociology at Indiana University.

■ **MICHAEL SWAFFORD** (Sex Differences in Soviet Earnings) is Assistant Professor of Sociology and Anthropology at Vanderbilt University. He is studying children's perceptions of occupations; his work is based on data gathered in the United States, Japan and the U.S.S.R.

■ **SANDRA SCARR** (Influence of "Family Background" on Intellectual Attainment) is Professor of Psychology at Yale University. **RICHARD A. WEINBERG** is Professor of Psychoeducational Studies at the University of Minnesota. Together they have worked on genetic and environmental transmission of intellectual, personality, and attitudinal characteristics in families. Their research has included studies of the effects of transracial adoption on the development of black children. Their book, *Child Development* (Harcourt, Brace, Jovanovich), is forthcoming.

■ **ISRAEL ADLER** (Cross Pressures during Socialization for Medicine) is Lecturer at the Hebrew University of Jerusalem. He is interested in social stratification and methods of analysis of longitudinal data. **JUDITH T. SHUVAL** is Associate Professor of Sociology at the Hebrew University of Jerusalem. Her work is on the socialization of health professionals, the processes of adaptation of immigrant physicians from the Soviet Union in Israel, and health behavior. She is author of *Social Functions of Medical Practice* (Jossey-Bass, 1970) and *Processes of Socialization for Health Professions* (forthcoming).

■ **E. M. BECK** (Stratification in a Dual Economy) is Associate Professor of Sociology at the University of Georgia. His current research is focused on the application of dual economy and labor market theories to understanding the interrelationships among discrimination and poverty and the organization of the industrial order. **PATRICK M. HORAN** is Associate Professor in the Department of Sociology at the University of Georgia. He is doing research (with Beck and Tolbert) on alternatives to neoclassical theory. He also is working on an analysis of the structure of scientific research areas. **CHARLES M. TOLBERT II** is a Ph.D. Candidate in the Department of Sociology at the University of Georgia. He is studying ways of identifying economic sectors empirically. His dissertation focuses on the implications of dual economy theory for research on occupational mobility.

■ **KAREN S. COOK** (Power, Equity, Commitment in Exchange Networks) is Assistant Professor of Sociology at the University of Washington. She is doing experimental research on the linkage between equity processes in exchange networks and the exercise of power and the formation of coalitions. **RICHARD M. EMERSON** is Professor of Sociology at the University of Washington. His research interests include laboratory research on power in exchange networks and field and historical research on exchange networks and authority structures in South Asian agrarian society.

■ **DAVID E. PAYNE** (Cross-National Diffusion: Effects of Canadian TV) is Associate Professor in the Department of Sociology at the University of North Dakota. He is conducting longitudinal research on the effects of Canadian TV on U.S. viewers and cross-sectional research on the effects of U.S. TV on Canadian viewers.

■ **NICHOLAS L. DANIGELIS** (Black Political Participation in the United States) is Assistant Professor of Sociology at the University of Vermont. He is continuing his historical analysis of black politics in the U.S. via survey and archival data.

■ **WAYNE J. VILLEMEZ** (Comment on Szymanski, ASR June, 1976) is Associate Professor of Sociology at the University of Illinois, Chicago Circle.

New Publications Available from ASA...

1978 Guide to Graduate Departments of Sociology

Identifies faculty by name, highest degree, where and when degree was earned, and areas of interest.

Members & Students: \$4.00

Non-members & Institutions: \$10.00

1978 ASA Directory of Members

Members: \$5.00

Non-members & Institutions: \$10.00

1978 Directory of Departments of Sociology

Provides name of department; chairperson or representative; department's mailing address and phone number; number of faculty in department; number of students; highest degree offered. 1,929 institutions listed.

Members: \$5.00

Non-members & Institutions: \$10.00

Journal of Health & Social Behavior Index

(Volumes 1-17, 1960-1976)

Members: \$5.00

Non-Members & Institutions: \$7.50

To order any of the above, check publication(s) desired and send coupon, enclosing payment, to: American Sociological Association, 1722 N Street, NW, Washington, DC 20036.

(All orders must be prepaid)

Name _____

Address _____

City

State

Zip

THE PRESIDENTIAL ADDRESS: CUMULATIVE CHANGE IN THEORY AND IN HISTORY*

AMOS H. HAWLEY

University of North Carolina

American Sociological Review 1978, Vol. 43 (December):787-796

A comprehensive overview of the literature on social change would almost certainly bewilder the uninitiated person should that individual have the patience to undertake so arduous a task. The observer would find that the spectrum ranges from Robert Nisbet's (1969:197) denial of change, there being only a "finely-graded, logically continuous series of 'stills,' as in a movie film. . .," to Wilbert Moore's (1963:11-2) assertion that change is ubiquitous. Running the gamut, it seems that change covers any difference between before and after states, regardless of the units affected, the magnitude of the differences, the time interval involved, or the repetitiveness of the difference. What is true in the aggregate, of course, is not true of individual scholars. Each has employed a definition designed for a particular purpose. I shall do likewise. By social change I mean any nonrecurrent alteration of a social system considered as a whole. The term *nonrecurrence* in the definition is intended to exclude rhythmic events, such as the waking-eating-sleeping round of the diurnal cycle, daily trips to and from work or school, the annual cycle of holiday festivities, the succession of generations, and other such pulsations. These are the means by which a given pattern of relationships is sustained rather than altered. A more difficult exclusion concerns

short-term variations around a central tendency, inasmuch as they usually are recognized as such only in retrospect. Try as we might, we have perfected no way of recognizing such variations for what they are at the time of their occurrence.

Nonrecurrent alterations appear in many forms. Of these the most significant, if least dramatic, is what may be called *cumulative change*. This may occur as a single increment to the content of a social system or it may be comprised of a series of increments, each of which prepares the way for the next. In either case it constitutes growth of the system, a movement from small and simple to large and complex.

Why, it is reasonable to enquire, should one expect to find growth or cumulative change in social systems? Various answers suggest themselves. One rests upon an analogy with change in biological organisms. That the organism begins with a fertilized cell which, under appropriate conditions, subdivides repeatedly to produce increasing size and structural elaboration is common knowledge. Whether the processes involved in that phenomenon have counterparts in the growth of the human social system is too debatable to occupy us here. A second kind of answer to the question of "why cumulative change?" may be found in an argument from history. The historical record enumerates many instances of the rise of empires and of civilizations from small and simple beginnings. The fact that many have declined and disappeared does not gainsay the tendency to cumulative

* Address all communications to: Amos H. Hawley; Department of Sociology; University of North Carolina; Chapel Hill, NC 27514.

Presented as the Presidential Address to the American Sociological Association at its annual meeting in San Francisco, September 1978.

change. It merely poses another problem. Lessons from history are most convincing, however, when they can be shown to conform to a pattern that can be stated as a set of principles. That, if it can be demonstrated, would constitute a third and most satisfactory answer to the question.

The issue, then, may be stated as: Is there a pattern in cumulative change? To pose the question in this manner may seem to minimize the part decision processes play in change. No one can deny that individuals calculate means to ends. But whether purposefulness in the individual has any necessary outcome in the aggregate is moot. In any event, I do not wish to be mired in an attempt to distinguish intended from unintended effects. The question to be considered here is simply: Are there kinds of events or circumstances which lead inexorably toward cumulative change? In my following remarks I shall treat this question by exploring the implications of three ideas current in much of our thinking, namely, irreversibility, evolution, and expansion.

Irreversibility

The notion of directionality in change rests upon the assumption of irreversibility. This assumption is also basic to a holistic or social system approach to the treatment of change. Irreversibility may be due, as A. J. Lotka (1924: chap. 3) has pointed out, to the mere improbability that elements, after having been moved about randomly, can be immediately restored to their original order. The experience of picking up a deck of cards that has been dropped to the floor illustrates the point. That may seem to have a parallel in instances in which a major catastrophe so disorders the relationships in a social system that they cannot be reconstituted in the prior pattern. Such a situation, however, bears a closer resemblance to an omelet than to a disordered deck of cards: in neither case can the new arrangement of substances be unscrambled. But there is a third circumstance producing irreversibility. That occurs with the creation or emergence of new properties when previously separated organic units are brought

into interaction with one another. The relationship thus formed is not inherent in any of the individuals.¹ While some relationships may be terminated, others may not, for interdependence is a survival imperative. The Hobbsian contract is not one that can be readily broken. Irreversibility is thus a condition of structural accumulation as well as of structural rearrangement. As applied to system change, irreversibility does not preclude the possibility of decline and even disappearance. It means rather that the succession of events by which a system was brought to a given state cannot be followed backward to a starting point. A different path must be followed in decline and it, too, moves through a nonreversible sequence.

Perhaps all of this is familiar enough, but an implication often overlooked, binding though it may be, is that the explanatory principles developed at one level of integration are not applicable to another level. To declare with Professor Homans (1964) that this conclusion is false because no satisfactory principles of explanation for higher levels of social integration have been demonstrated begs the question. What is at issue is how problems are stated, the selection of variables, and the modes of operationalization of variables. Nor has it been convincingly shown, contrary to the arguments of methodological individualists (cf. Watkins, 1953), that attitudes, perceptions, and other such conceptualizations are any more palpable than are relationships and combinations of relationships. No one, to my knowledge, has yet reassembled the whole human being from the many analytical abstractions employed in the pursuit of generalizations.

The reductionist is usually misled by a methodological tactic which the holist uses when engaged in quantification. To describe a population as consisting of so many individuals is no more a confession that the ultimate reality is the individual than is the measurement of a farm in acres

¹ "Whenever certain elements combine and thereby produce, by the fact of their combination, new phenomenon, it is plain that these new phenomena reside not in the original elements but in the totality formed by their union." (Durkheim, 1938; pp. xlvii)

of land an admission that only an acre is real. Similarly, while a single birth is an individual experience, a birth rate is a structural feature, and the explanations of the two facts may have little in common. The aggregate, the reductionist's conception of social reality, may be just an aggregate if composed of units thrown together simply because they conform to a given definition. But, if the aggregate is not the product of a statistician's convenience, then an explanation of its existence must lie elsewhere. Pursuit of the *elsewhere* takes one to a social structure.

It seems, then, that a theory of social change cannot be designed to explain both individual variations and social system variations, as Gudmund Hernés (1976) would have it. The individual life cycle has no counterpart in the duration of a social system. Functional positions can remain constant despite the turnover of incumbents. That the converse is true seems very unlikely. Change in a social system alters the life conditions of all participants and they must adapt in order to remain in the system. Irreversibility is the bridge linking levels of integration, the similarities among which lie mainly, if not exclusively, in analogy.

Evolution

Irreversibility is an elemental assumption in the theory of evolution.² Whether that concept is transferable from biotic to social levels of analysis depends a great deal on how it is interpreted. Certainly attempts to draw very close parallels between levels risks falling into the reductionist fallacy. The principle of irreversibility tells us there can be no general law of evolution operating across all system levels (Jacob, 1977). The properties distinguishing one level of integration serve as postulates for the next higher level. But the hypotheses designed to account for phenomena at the higher level must be cast in terms of the peculiar constraints that operate in and upon that level. Yet there are two respects in which the evolution concept has a generic meaning. Sub-

stantively, it implies a change from simple to complex forms.³ Analytically, change is viewed as proceeding through variation of units and natural or fortuitous selection. A necessary caveat in this latter respect is that the selection hypothesis is a tautology, albeit a highly useful one; but like all tautologies it suffers from a two-way causation. A point too little stressed is that societal evolution appears to be a Lamarckian, rather than a Darwinian phenomenon.

One of the characteristics of traditional evolution thinking is the notion of discontinuity in change, or of change as moving through a stage-like progression. The simplicity of the stage concept is beguiling, despite the precariousness of the implied equilibrium assumption. It is doubtful, however, that an equilibrium assumption can be escaped, not even by those who argue that change is imminent, for change can only be imminent when there is not-change. Equilibrium waits in the wings of the mind to move to center stage at every recognition of unit character. And, without the supposition of unit character, order in the universe would not be conceivable. More to the point of this discussion is the need for some means whereby historical time can be converted to analytical time. The stage concept serves that purpose for better or for worse. Despite its many detractors, the concept has shown remarkable vitality. It is an essential ingredient in the bioecologist's concept of the ecosystem, which has its emulators in social science, and it lives on in the social evolutionist's chronology of societal types.

It is noteworthy that biologists have turned from typological to probability models when the concern is with the phylogenetic problem, but they employ an equilibrium model when the ecological problem arises. The difference is more than a matter of preference; it points to a distinction between evolution and growth. Evolution deals with the appearance of new species or forms of whatever kind,

² A more fundamental application of the term occurs in astrophysics. See Benjamin Gal-Or (1972).

³ G. E. Swanson (1971:3) uses the term to mean "capacity to exploit environment." This usage verges closely upon the meaning of progress. I prefer a more neutral conception.

while growth pertains to the maturation of a form to a point at which, presumably, the form is capable of reproducing progeny which are then subject to natural selection. In contrast to the variation-selection model of evolution, the growth model is best represented by the logistic curve. There is much less consensus on what takes place in the passage along the curve. If we can resolve that difficulty, we may be in a position to determine how the two concepts accord with the history of societal development.

Expansion

What is meant by cumulative change requires a definition of a social system. That term, as used here, denotes an arrangement of routine activities or roles and relationships by which a population sustains itself in a given environment. Cumulative change, then, refers to increase in the number and variety of roles and relationships. Such increase, however, cannot proceed far without increases in population and in territory. Population increments are needed, not just for the augmented variation it brings, but also to staff the growing structure of roles and relationships. Added population is needed also to provide a clientele, i.e., a market, for any increases in productivity that might occur. Differentiation without a market is like a building without a foundation. Increases in structural complexity also make demands for access to an enlarging territory from which to obtain food and raw materials and in which to find room for the conduct of activities. Complexity and scale are intimately linked, though the linkage may not always be superficially evident. Hence the use of the term expansion to characterize societal growth.

So far as I am aware, Lenski (Lenski and Lenski, 1978) is one of the few students of evolution to give even casual reference to expansion. Dudley Duncan (1964; see also Eisenstadt, 1964), who shares the interest in evolution, suggested that expansion may be the key to the transition from one major level of social evolution to another. But very little systematic work on the concept has been forthcom-

ing. The term has been commonplace in the historical literature for many years, but it has been used merely to describe particular series of events, instances of which crowd the historical record (Gras, 1922; Woodruff, 1966). Braudel (1966:660) speaks of the "logical laws of expansion," but I have been unable to find them in his work. Occasional efforts to generalize the historical experience have been made, but those have been incomplete (McKenzie, 1934; Hawley, 1950: chaps. 18, 19, 20; 1971; Gutkind, 1953; Quigley, 1961). Empirical studies, of which there have been many, have dealt mainly with events at the margins of expanding systems (Wilson and Wilson, 1954; Gough, 1955; Rao, 1970; Vidich and Bensman, 1960). More recently, the growth of interest in world systems promises a use of historical materials in a systematic development of the expansion concept (Wallenstein, 1974; Choucri and North, 1974). Very little of this work has found evolution theory a helpful basis from which to proceed. That may be because a point is reached in the growth of social systems beyond which evolution theory is no longer helpful. I will want to return to this suggestion later in my remarks.

Whether the process of cumulative change leads to evolution or growth depends, other circumstances constant for the moment, on how concurrent are advances in complexity and scale. There have been many instances in the past in which population has increased significantly without corresponding increases in structures, as a result of reductions of enemies or of windfalls in the food supply. If the loss of proportion between the population carrying capacity of the system and the number of people on hand is the only disturbance that has occurred, the effect is a budding off of colonies. The colonies move off in search of niches in the environment in which they might settle. They may be likened to progeny possessed of a range of variability and subject, therefore, to selection by environment. Some survive, and in doing so adapt their structures to new circumstances. Others succumb as a result of their inability to come to terms with unfamiliar environments. In this way evolution of social systems may be con-

ceived. That is, through a combination of happenstance and adaptive success one or more complex or advanced social systems are produced (cf. Simpson, 1967; Jacob, 1977).

On the other hand, when complexity and scale advance more or less together, the effect is growth or expansion rather than evolution. But now it is necessary to enquire into how that process is begun, for that it can have a spontaneous causation is most unlikely. The axiom that a thing cannot cause itself is as applicable here as elsewhere.

The normal, if not the necessary, condition for expansion arises from the colonization process described in the preceding paragraph. A spread of settlements over an area, under appropriate conditions, may create a social field, a universe of more or less frequent interactions among the settlement units (cf. Lesser, 1961; Wilkinson, 1970). A field may be visualized as a territory over which the several settlements, or centers, each with a tributary area differing in scope and resource composition, are variously linked in a common transportation network. One or two of the centers is situated at the intersection of intra- and interregional routes of travel, a much larger number are located at the crossings of internal thoroughfares, and some are found at the extremities of interior routes. The field notion, it will be noticed, substitutes an assumption of interdependence of units for the assumption of independence in evolution theory.⁴

Social fields are a commonplace in historical experience. History describes a western succession of such interaction networks. One of the earliest recorded centered upon Babylon in the Euphrates valley. It was followed by a field of greater dynamics in the Mediterranean region where Miletus, Athens, Alexandria, Rome, and Constantinople served successively as focal points. Still later, ascendancy shifted to northwestern Europe from which subsequently the interaction

network was extended more and more widely.

Within an interaction field, disequilibrium, which is a requisite for system change, is a chronic condition.⁵ Not only are there disturbances arising from the biophysical environment, but also each center is exposed, unequally to be sure, to repeated challenges from the social environment. Travellers circulate through the network, bearing ideas, experiences, and artifacts, that is, information and misinformation, originating from local and extralocal sources. Information piles up, as it were, in the most accessible center and drifts outward to the less favorably located centers. Of all the kinds of information that flow into and through the network, those which affect facility in movement are doubtlessly the most critical. As Joel Smith (1968) has pointed out, they determine the measure of accessibility, set the terms of expansion, and ultimately, as we shall see, impose limits to growth. If invention is a drama enacted on a crowded stage, as Michael Polanyi (1959:117) has said, there must be means for bringing actors with diverse experiences together. And as inventions enter into use they call new relationships into being, every one of which involves a transportation of some kind. Technical accumulation begins in mobility and is sustained by mobility.

As a general proposition it may be stated that the complexity and scale of a social system are a joint function of the efficiency of its techniques for the movement of people, materials, and information. Efficiency is measured most cogently by the number of activities that can be articulated per unit of time and cost. It derives partly from the tools for movement which, together with the knowledge for their fabrication and use, comprise what is ordinarily regarded as technology. But efficiency rests also on the less obvious, though no less important, organizational arrangements essential to the application of the tools. The relays of messengers employed for the integration of ancient empires was an organizational

⁴ This distinction also has been recognized by Donald Campbell (1965) in his paper on "Variation and Selective Retention in Socio-Cultural Evolution."

⁵ According to Göster Carlsson (1968), the rate of change varies with the extent of disequilibrium.

device. So also are the freight forwarding agency, the commercial bank, and the insurance firm of a much later period. A more subtle number of this class is standardization of the terms of discourse, including language to be sure, but also weights and measures, coinage, units of time, rules of the road, standards of judgment, and, above all, forms of organization. The technology for movement is in no sense peculiar in its composition. Tools and organization are two sides of the same coin. Technology is nothing more nor less than the instrumental aspect of culture (Boulding, 1969).

These distinguishable components of technology are often staggered in their development. In many instances it seems that the tool appears before an effective organization for its use has been devised and that happens before the many adjunctive behaviors have made their accommodations. Alfred Chandler (1977; see also Taylor and Neu, 1956) describes how experience with the administration of the railroad had to accumulate before a solution was found in an hierarchical management structure some twenty years (in the 1860s) after the steam railway was introduced. But another twenty years passed before a standardization of time and of track gauge were accomplished, in the 1880s. One of the more time consuming phases of standardization, especially where relationships are bridging cultural differences, is with reference to forms of organization and the procedures by which they operate. For that requires a resocialization of not just the few users of the imported tools but of large sectors of a population and eventually of the entire population. In the end the effect will occur if the relationship is uninterrupted and if it has gained vital significance.

The tool-organization-standardization sequence creates conditions out of which other, different sequences unfold. For example, the effects of standardization as a facilitator of movement are not confined to one range of events. The standardization of railroad track gauge in the United States opened wider the gates to interregional flows of information and, thus, increased opportunities for invention. Similarly, an organizational form developed in

connection with the application of one mechanical contrivance is often transferable to others. Chandler (1977: chap. 9) makes a point of how the management structure devised for the dispersed operations of the railroad became a model for other large-scale enterprises. An exogenous influence rarely ends with a single response in a system; rather does it produce a concatenation of effects that terminates only when the system has completed its absorption of the new element. The process, comprising numerous feedback loops, gives to cumulative change a helical pattern of progression.

In the context of the interactive field the cumulative process is most rapid in the strategically located center. Accordingly, it gains an increasing capacity to mediate and coordinate a diversity of activities scattered over a widening area. The expansion process, as McKenzie (1934) observed some time ago, involves countervailing currents of redistribution. A centripetal movement of selected specialists and ancillary workers parallels the centralization of information to develop the institutions through which a center grows in size and in administrative power. A centrifugal movement of explorers, raw material extractors, processors, and managers carries technical acquisitions into resource developments on a receding frontier. Thus the field becomes organized in a hierarchy of centers and local tributary zones with the result that a single expanded system replaces a number of localized and relatively independent systems.

This is a process that can work on any scale, limited only by the technology for movement. As a matter of fact, however, it generates what might be regarded as organizational equivalents of scale. On the one hand, it radically changes the significance of population in local areas. The merging of lesser units into a larger unit is also a pooling of demographic resources. Each subsystem gains the benefit of a much larger labor force and range of skills than it could muster within its original boundaries. This is by way of compensation for the loss of local autonomy. On the other hand, expansion alters the orientation to territory. In the early phases

communications are hampered by organizational differences between center and periphery. In consequence expansion can only proceed by the imposition of the center's institutional forms and procedures upon outlying settlements and that called for political domination of territory. But in the long run sustained interaction proves an equally, if not more effective, and a much less costly, means of producing structural convergence than political coercion. Political domination gives way before an increasing ease of access to territory.

It is true, of course, that in the past processes of expansion have run their course and come to an end many times over. Various explanations of the conclusion of expansion phases have been proposed. Ralph Turner (1941: Vol. 2, 1298), the historian, contended that a system can reach only to the outer edge of the region to which its agricultural techniques are adapted. Neither colonization nor the supply of armies is possible on a sustained basis beyond that limit. Carroll Quigley (1961) finds the end of expansion in the natural decay that sets in when the uses of surplus wealth become institutionalized. As that occurs energies are directed from invention to maintenance of bureaucratic positions. A case in point might be in the institutionalization of slavery in the Roman Empire which, according to F. W. Walbank (1946:19), so impoverished the citizen population that the local market contracted and invention lost its incentives. The sinologist, Mark Elvin, finds a different cause of the termination of expansion. He says that "empires tend to expand to the point at which their technological superiority over their neighbors is approximately counterbalanced by the burdens of size" (Elvin, 1973:19).

Each of these statements describes a conception of how a return to equilibrium comes about. The unspoken assumption in each account is that the possibilities for growth are contained in and limited by a given technology for movement. The limits are approached by virtue of the difference between the exponents of increase in scale and complexity. While numbers of

people and of activities increase additively, relationships increase by multiplication. In consequence a rapidly mounting density of interaction generates steeply rising costs of movement of people, goods, and messages. A point is reached in the density curve beyond which the costs of movement are too high to support any further elaborations of structure. Any tendency to exceed that density, *ceteris paribus*, results in a return to scale, as Boulding (1953) has observed in his seminal paper on a theory of growth. The return to scale in simple systems is manifested in the colonization process mentioned earlier. In more complex systems it may develop as a strain toward decentralization of authority, forming subsystems, and restoring limited degrees of autonomy to local groups.

The growth process can always be resumed, of course, when further improvements in the technology for movement are introduced. Since that is apt to occur in some systems before it affects others, the former tend to expand into the territories of the latter, thereby absorbing them into an enlarging system. Several major mobility revolutions have subdivided Western history into expansion phases. Earliest among these was the transition from animal powered to mechanically powered movement. That began with the reliance upon wind and sail to venture beyond coastal waters and reached a maximum scope with the application of steam power to over-land as well as to over-water movements. In that regime it was possible to extend systems far beyond localities and harness the resources and markets of distant peripheries to center of expansion. It has been suggested that the potentialities of the steam power era were fully utilized by 1870 and that very little change took place in the next several decades (Landers, 1966). But that is to overlook the telegraph, which appeared in the 1840s and launched a second important transition; namely, a separation of communication from transportation. By this means a new dimension was added to the territorial organization of economic, political, and other activities. For the first time central

offices could exercise control on a daily basis over widely scattered branch offices and producing sites. Close upon the heels of the communication revolution came dramatic improvements in the facility for short-distance movement. The telephone and the electric street railway in the third quarter of the nineteenth century and the motor vehicle at the turn of the century gave a new scope to the pattern of local relations which had received only minor alterations since the domestication of the horse. The last of the major turning points has been in the making for some time and is not yet concluded. This is an advancing substitution of communication for transportation. It was initiated with the telegraph and telephone, carried further by radio, radar, and television and has reached a present apogee in satellite and laser beam transmissions. Technical information, management instructions, credit and foreign exchange can now move independently of transportation. Moreover, the reduction of time distances to near zero have all but eliminated the boundaries of systems.

An inference from the hypotheses which has guided this discussion is that efficiency in movement is an accurate indicator of the extent of cumulative change. That admittedly has not been fully demonstrated. There is, however, a mounting volume of evidence to support the argument that a measure of mobility can serve that purpose. Wilfred Owen (1964:14) has shown that freight and passenger mobility are sensitive indicators of the level of economic development. Numerous correlation studies of comparative data reveal measures of communication to be among the most indicative of the level of other measures of development (United Nations, 1968; Olsen, 1968; Cutright, 1963). And Henry Barbera (1978), in an unpublished paper, has effectively predicted the extent of external relation from the assumption that the power of a system is measured by the degree to which internal communications are developed. Mobility measures rival energy production as indicators of complexity or growth. But they are not qualitatively equal. Energy produced and consumed may be highly

concentrated in a few localities or economic sectors. Transportation and communication, on the other hand, constitute a more distributed feature of a system.

Conclusion

In conclusion, it appears that the course of history has progressively reduced the utility of an evolution model in the explanation of cumulative change. Although social variation as represented in occupational and territorial specialization has been carried to an unprecedented degree, it does not satisfy the assumption of the model; instead of the required capacity for independence of action, the differentiated parts are inescapably caught up in a tightening web of interdependences. This does not mean an end to, or even a slowing of, change. The accumulation of information in storage facilities—a means of transportation through time—is so vast that generations will pass before its potential uses are exhausted. In the meantime the impetus to change may be expected to shift from center to center, further reduction of cultural differences will occur, and the density of organization on a world scale will continue to increase. But the disappearance of evolution as a mode of change in human social systems, if true, is not without risks. Whereas in the anarchy of a multiplicity of localized systems fatal errors could be made here or there without jeopardizing the survival chances of other systems, that can no longer be expected. The single world system has a limited tolerance for error. It must either acquire methods of anticipating and compensating for significant errors or fail to survive. In that event the evolution process may once again take command of the course of change. The opportunities for speculation on this theme are inviting, but space does not permit me to pursue them.

REFERENCES

- Barbera, Henry
1978 "The international mosaic: contact, conflict and cooperation in world society." Unpublished paper.

- Boulding, Kenneth
1953 "Toward a general theory of growth." *Canadian Journal of Economics and Political Science* 19:326-40.
1969 "Technology and the changing social order." Pp. 126-40 in David Popenoe (ed.), *The Urban Industrial Frontier*. New Brunswick: Rutgers University Press.
- Braudel, Fernand
1966 *The Mediterranean and the Mediterranean World in the Age of Philip II*. New York: Harper and Row.
- Campbell, Donald
1965 "Variation and selective retention in socio-cultural evolution." P. 30 in H. Barringer, G. I. Blankston, and Ray Mack (eds.), *Social Change in Developing Areas*. Cambridge, Ma.: Schlenkman.
- Carlsson, Göster
1968 "Change, growth and irreversibility." *American Journal of Sociology* 73:706-14.
- Chandler, Alfred D.
1977 *The Visible Hand: The Managerial Revolution in American Business*. Cambridge, Ma.: Harvard University Press.
- Choucri, Nazli and Robert C. North
1974 *Nations in Conflict: National Growth and International Conflict*. San Francisco: W. H. Freeman.
- Cutright, Phillips
1963 "National political development." *American Sociological Review* 28:253-64.
- Duncan, Otis D.
1964 "Social organization and the eco-system." P. 57 in R.E.L. Faris (ed.), *Handbook of Modern Sociology*. Chicago: Rand-McNally.
- Durkheim, Emile
1938 *The Rules of Sociological Method*. Glencoe: Free Press.
- Eisenstadt, S. N.
1964 "Social change, differentiation and evolution." *American Sociological Review* 29:375.
- Elvin, Mark
1973 *The Pattern of the Chinese Past*. Stanford: Stanford University Press.
- Gal-Or, Benjamin
1972 "The crisis about the origin of irreversibility and time anisotropy." *Science* 176:11-7.
- Gough, Katherine
1955 "The social structure of a Tanjore village." Pp. 36-52 in McKim Marriott (ed.), *Village India*. Chicago: University of Chicago Press.
- Gras, N.S.B.
1922 *An Introduction to Economic History*. New York: Harper Brothers.
- Gutkind, Frederick
1953 *The Expanding Environment: The End of Cities and the Rise of Communities*. London: Freedom Press.
- Hawley, Amos H.
1950 *Human Ecology*. New York: Ronald Press.
1971 *Urban Society: An Ecological Approach*. New York: Ronald Press.
- Hernes, Gudmund
1976 "Structural changes in social processes." *American Journal of Sociology* 82:513-47.
- Homans, George C.
1964 "Bringing man back in." *American Sociological Review* 29:809-18.
- Jacob, Francois
1977 "Evolution and tinkering." *Science* 196:1161-6.
- Landers, David
1966 "Technological change and development in Western Europe, 1750-1914." P. 464 in M. M. Postan and J. J. Babakkuk (eds.), *The Cambridge Economic History of Europe*. Cambridge, Eng. The University Press.
- Lenski, Gerhard E. and Jean Lenski
1978 *Human Societies: A Macrolevel Introduction to Sociology*. 3rd ed. New York: McGraw-Hill.
- Lesser, Alexander
1961 "Social fields and the evolution of society." *Southwestern Journal of Anthropology* 17:40-8.
- Lotka, Alfred J.
1924 *Elements of Mathematical Biology*. Baltimore: Wilkins and Wilkins.
- McKenzie, Roderick D.
1934 "Industrial expansion and the interrelations of peoples." Chap. 2 in E. B. Reuter (ed.), *Race and Culture Contacts*. New York: McGraw-Hill.
- Moore, Wilbert E.
1963 *Social Change*. Englewood Cliffs: Prentice Hall.
- Nisbet, Robert A.
1969 *Social Change and History*. London: Oxford University Press.
- Olsen, Marvin
1968 "Multivariate analysis of national political development." *American Sociological Review* 33:699-712.
- Owen, Wilfred
1964 *Strategy for Mobility*. Washington, D.C.: Brookings Institution.
- Polanyi, Michael
1959 *The Study of Man*. London: Routledge and Kegan Paul.
- Quigley, Carroll
1961 *The Evolution of Civilizations*. New York: Macmillan.
- Rao, M.S.A.
1970 *Urbanization and Social Change: A Study of a Rural Community on a Metropolitan Fringe*. New Delhi: Orient Longman's.
- Simpson, G. G.
1967 *The Meaning of Evolution*. New Haven: Yale University Press.
- Smith, Joel
1968 "Transportation: social aspects." Pp. 128-34 in *International Encyclopedia of the Social Sciences*, Vol. 16. New York: Macmillan.
- Swanson, G. E.
1971 *Social Change*. Glenview: Scott, Foresman.
- Taylor, George R. and Irene O. Neu
1956 *The American Railroad Network: 1861-*

1890. Cambridge, Ma.: Harvard University Press.
- Turner, Ralph
1941 *The Great Cultural Traditions*. New York: McGraw-Hill.
- United Nations
1968 "Urbanization: development policies and planning." *International Social Development Review* 1:30.
- Vidich, Arthur and Joseph Bensman
1960 *Small Town and Mass Society*. Garden City: Doubleday.
- Walbank, F. W.
1964 *The Decline of the Roman Empire in the West*. London: Cobbett Press.
- Wallenstein, Immanuel
1974 *The Modern World System*. New York: Academic Press.
- Watkins, J. W. H.
1953 "Ideal types and historical explanation." Pp. 723-43 in Herbert Fiegl and May Brodbeck (eds.), *Readings in the Philosophy of Science*. New York: Appleton-Century-Crofts.
- Wilkinson, Kenneth P.
1970 "The community as a social field." *Social Forces* 48:311-22.
- Wilson, Geoffrey and Monica Wilson
1954 *The Analysis of Social Change*. Cambridge, Ma.: The University Press.
- Woodruff, William
1966 *Impact of Western Man: A Study of Europe's Role in the World Economy, 1950-1960*. New York: St. Martin's Press.

MANUSCRIPTS FOR THE ASA ROSE SOCIOLOGY SERIES

Manuscripts (100 to 300 typed pages) are solicited for publication in the *ASA Arnold and Caroline Rose Monograph Series*. The Series welcomes a variety of types of sociological work—qualitative or quantitative empirical studies, and theoretical or methodological treatises. An author should submit three copies of a manuscript for consideration to the Series Editor, Professor Robin M. Williams, Jr., Department of Sociology, Cornell University, Ithaca, New York 14853.

THE IDEOLOGY OF CHILDHOOD AND THE STATE: RULES DISTINGUISHING CHILDREN IN NATIONAL CONSTITUTIONS, 1870-1970*

JOHN BOLI-BENNETT
University of Santa Clara

JOHN W. MEYER
Stanford University

American Sociological Review 1978, Vol. 43 (December):797-812

In the modern world the social role of childhood is increasingly differentiated from adulthood and brought under the jurisdiction of the state. This is not primarily a social organizational process but an institutional or ideological one: the ideological rules of differentiated and state-managed childhood diffuse throughout the global nation-state system regardless of variations among nations in dimensions of organizational development. We show this by coding the constitutions of all nation-states since 1870 for references and rules concerning childhood. Our index summarizing these references rises sharply between 1870 and 1970. At any point, the index does not vary significantly among rich and poor, new and old, or central and peripheral states. Ideological rules concerning childhood are weakly (and decreasingly) related to technical development, and strongly (and increasingly) related to the general authority of the state.

In modern societies children's roles become increasingly distinct from those of adults (Ariès, 1962; de Mause, 1974). In everyday interaction children are segregated from many aspects of adult life. Organizationally, elaborate structures creating separate positions for children arise: schools, prisons, libraries, medical facilities, and so on. At the same time childhood becomes distinct in institutional rules and ideology: new rules and language develop, locating young persons in separate categories and defining special rights and responsibilities for them (Ariès, 1962; Skolnick, 1975). All these changes occur under the aegis of the nation-state, which plays an increasing role in the regulation of childhood.

So far as the welfare of children is concerned, these changes evoke the most varied interpretations. Ariès calls attention to the dependence, repression, and formal inequality involved in the insulation of children from adult society, while others

see the same phenomena as offering more protection and equality (e.g., de Mause). Both elements are clearly involved: modern childhood is both more dependent and more protected than childhood in the past. In one sense children are less equal to adults than they used to be; in another sense they are more equal.

In this paper we trace the emergence of differentiated and state-managed childhood at the institutional level over the last century; we focus on the fundamental rules of nation-states expressed in national constitutions. Such rules embody very general ideological definitions of the nature of persons and society. In the short run they may have little to do with organizational reality. We argue, however, that the expanding institutional structure of differentiated childhood diffuses around the world precisely as ideology, rather than as a direct reflection of social organization. The rapidity of this diffusion arises from the increasing integration of the world stratification system and the relative homogeneity of its standards of value. We suppose that, in the longer run, the establishment of the ideology of modern childhood throughout the world lends strong support to the organizational implementation of dependent but protected childhood status.

Two driving forces, discussed in more

* Address all communications to: John Boli-Bennett; Department of Sociology; Stanford University; Stanford, CA 94305.

The research reported here was conducted at, and supported by funds from, the Boy's Town Center for Youth Development, Stanford University. However, the opinions expressed or the policies advocated herein do not necessarily reflect those of Boy's Town.

detail below, underlie our perspective. First, the technical rationalization of the modern world system defines individuals as central social units and rationalizes individual life and action (Weber, 1930). The purposive socialization of children in a differentiated childhood status appears crucial for the rational collective production of capable, loyal, and responsible individuals. Second, as the stratification system and culture of rationalization become both worldwide and ideologically compelling, national states greatly expand their corporate authority (Swanson, 1971; Meyer et al., 1975). They increasingly take formal jurisdiction over the management of key aspects of rationalized collective life, including differentiated childhood.

BACKGROUND

Most theories of the differentiation of childhood implicitly assume that institutional rules directly reflect changes in social organization. Conceptions of childhood derive from networks of child-oriented organizations and the immediate individual and group interests these networks embody (Meyer and Nagel, 1975). Consider demographic arguments: the decline in the birth rate makes individual children more valuable and makes more resources available for each child; the decline in childhood mortality makes social investments in specific children more profitable. Consider economic arguments: skill and socialization requirements rise with economic growth; more resources are available with growth; development produces unemployed youth and engenders mechanisms to control these youth. Finally, consider familial and social arguments: nuclear families with parents working away from the home are less able to manage and socialize children; urbanization frees children from community and family control and increases problems of child management.

As in much of present sociological reasoning, these arguments view society-wide institutions as the aggregate reflections of interpersonal and local organizational change. They lead to the following general hypothesis:

The creation, in state and society, of general institutional rules differentiating childhood reflects the level of economic development and social differentiation prevailing in society.

This hypothesis seems plausible, but it does not account for the *rule-like* aspect of childhood in modern technicized societies. Childhood is more than a collection of schools, specialized professions, children's books and records, and separate courts. It is also a collection of institutional rules, or ideological principles, defining children as persons quite distinct from adults (see Skolnick, 1975). For instance, there are rules prohibiting child labor, requiring schooling, defining separate medical rights, creating special systems of criminal justice, protecting children from special types of assault and from inappropriate cultural materials, and defining special requirements for equality. All of these, whether implemented or not, constitute a special social status for children.

Do these general rules become institutionalized in society (or in the constitutive laws of nation-states) simply as a reflection of ongoing social organization? Many other organizational features of technical societies do not become so deeply entrenched in fundamental ideological doctrines. Ideologies have sources other than concrete organizational practices, and they affect—as much as reflect—organizational structures in society. We therefore propose an alternative general hypothesis:

The institutional rules defining childhood parallel the spread of the institutional system of technically advanced society, whose central components consist of the nation-state and its citizenship variant of individualism. As the technical system expands, childhood ideology becomes worldwide, associated with state authority rather than with variations in local social and economic organization.

ARGUMENT

General rules differentiating childhood and incorporating it as a distinct status in the structure of the state arise as part of the "modern" (bourgeois and technical)

ideology defining the nature of life and action. Two central elements of this perspective are relevant: the *nature of individualism* and the *expansion of the state*.¹

Individualism. In technically developed societies (and the technicized world-system) legitimate action is that undertaken by responsible and rational individuals rather than by families and other primordial or corporate groups. Successful economies and polities, it is understood, consist of effective individuals acting rationally. This vision characterizes both the socialist and capitalist variants of technical society, though societies differ in the way in which they conceive effective individual actions to be knit together in a collective economy and policy.

Life thus becomes *individual* life and biography, not the histories or traditions of corporate groups. But with the decay of primordially-based action, from what sources does the individual acquire the capacity to undertake effective choice and action? To some extent, the sources are conceived as biologically rooted, but biological theories conflict with ideologies of freedom and progress. The more important sources are those identified by socialization ideologies which claim that individual action is rooted in a socially-determined self. Contemporary institutional rules of childhood clearly assume a theory of socialization. They conceive of childhood as a period in which biological and social forces interact to generate the competent and effective *person* (individual). This conception justifies schooling, the differentiated culture of childhood, distinct patterns of legal and medical treatment, and so on.

The expansion of the state in the competitive world system. The preceding analysis helps explain the differentiation of childhood from other stages of biography; we next consider why the socialization of children has become such a prominent concern of the state.

Nineteenth century liberalism could see progress as the result of action by free individuals in a freely organized society.

¹ Our ideas here derive from a long tradition, including the work of Peter Berger (Berger et al., 1973), Jacques Ellul (1967; 1975), T. H. Marshall (1948), and Reinhard Bendix (1964).

In the twentieth century, the procedures and goals of progress have become clearly defined, throughout the global stratification system, as properties of collectivities (nations) operating as integrated organizational units. Thus, "progress and growth" (that is, the ideology of technical development) can, and even must, be pursued by the explicit action of the dominant collective social organizations: states. The models and doctrines of progress have become socially institutionalized tools of national development and national competition in a highly demanding world system. The emphasis on national success in the world system, giving great impetus to the expansion of states, increases their power and rapidly expands their jurisdiction over social life (Boli-Bennett, 1976; Meyer et al., 1975).

The state has the authority and the responsibility to hasten modern technical development, which is understood to depend heavily on the expanded productivity of *individuals*. The state must therefore socialize citizens to create a new "model of man" (Inkeles and Smith, 1974; consider the versions proposed by Mao, Nyerere, Castro—and B.F. Skinner). The state institutionalizes differentiated childhood as part of its responsibility to manage society's progress in terms of the standards of the world stratification system.² It assumes jurisdictional control of the socialization process (Tyack, 1966; 1974; Meyer et al., 1977). The ideology of differentiated childhood thus penetrates everywhere that the state's authority is on the rise, regardless of variations in local social organization.³

² This same argument helps explain state action to differentiate and manage other stages of individual biography, in particular that of the elderly (quite pointedly known as *senior citizens*). In the USA and elsewhere we commonly speak of the "national disgrace" represented by our treatment of the aged, and there is general ideological agreement that the state should act to redress this situation.

³ We consider as secondary the main issue in the literature—the question of how much autonomy and how many rights children should have (consider Ariès, 1962; Holt, 1973; de Mause, 1974). Individualism unaccompanied by state expansion tends to create curritanical orientations toward socialization because moral responsibility rests on the socialized individual. With the expansion of the state, individuals are seen as its agents (Bendix,

Diffusion of the ideology. Differentiated and state-managed childhood diffuses even more rapidly throughout the world at the institutional level than at the level of social organizational structure. It takes resources and time for poor countries to build schools, train teachers, and socialize families to hand over what were previously "their own" children to state organizations. The spread of the ideology of state-managed childhood, however, has fewer such impediments. Given a relatively integrated world economy, stratification system, and culture (Inkeles and Smith, 1974; Boli-Bennett, 1976), the diffusion of the general ideological depiction of modern childhood can proceed rapidly, like other formal aspects of the nation-state system. It may even penetrate more completely in the periphery than in the core: the ideology becomes crucial as a mobilization mechanism for the periphery trying to catch up with the core, from which the ideology originates. Also, the periphery tends in general to adopt ideology more expansively than the core because ideological development is much easier to institute than organizational change, and peripheral states try at least to give the strong impression that they are mobilizing in the accepted way.

THE RESEARCH PROBLEM

We explore the rise of state-generated differentiation of childhood by examining essential documents expressing contemporary ideology, national constitutions.⁴

1964). They are less autonomously responsible, and their natural impulses in childhood are less threatening. But both individualism and the rise of the state increase the institutional differentiation of childhood. This is the central historical process; varying attitudes toward child masturbation or children's political rights matter much less.

In this connection, note the association in the United States between the children's rights movement and demands for supportive state action (cf. Gross and Gross, 1977). This movement proposes an even more distinct status for children, but a status accompanied by more rights to be protected by the state. The movement is part of the general shift in American society away from puritan individualism toward a more expanded state.

⁴ Constitutions do not fully capture the extent of state involvement in childhood. In some instances constitutions understate state involvement because

We have studied all of the constitutions of independent nation-states for the period 1870–1970.⁵ The constitutions in effect at 20-year intervals (1870, 1890, 1910, 1930, 1950, 1970) were coded for references to childhood as a distinct stage in biography, and for references to state action in relation to children. Constitutions change rather frequently. Our work chronicles these changes, as well as the consequences of the creation of many new states in this period.

Following the discussion above, we hypothesize, first, that the increasing incorporation of childhood in the structure of the state characterizes all types of countries, rich or poor, new or old, central or peripheral. To study this we analyze changes in an aggregate index of constitutional provisions concerning childhood, discussed below as the *Childhood Index*. If any differences among types of countries are to be found, the diffusion idea discussed above leads us to anticipate higher scores on this index for peripheral countries than for the core countries.

Second, we assert that childhood as political ideology reflects not merely the general expansion of the world technical system, but the degree of authority over society assumed by the state. This proposition is evaluated by examining correlations between the Childhood Index and measures of the general authority of the state in society.

Third, because the rules defining childhood diffuse around the world as ideologi-

they contain a strong elastic clause that justifies state control over many areas not specified explicitly in the documents. More generally, much state involvement in childhood occurs through legislation, and a more complete analysis would consider legislation as well as constitutions. However, legislation reflects many factors besides the ideological processes discussed here, and a complex analysis would be required. Few researchers have yet tried to tackle legislation on a systematic comparative basis; for one effort in this direction, focusing on education, see Ramirez and Robinson (1979).

⁵ A complete list of sources of constitutions is available in Boli-Bennett (1976). A number of more or less comprehensive collections have appeared sporadically since about 1870, but many supplementary sources were used as well. Major collections, other than those mentioned in the text, include the works by Dareste and his successors (1891; 1910; 1928–34), Arosemena (1870; 1888), Posener (1909), and Peaslee's (1950; 1956) first two editions.

cal themes, they should be little affected by variations in the level of organizational complexity of national societies. We therefore expect low correlations between the Childhood Index and measures of economic and social development.

Finally, we hypothesize that the rules of differentiated childhood become *increasingly* closely tied to the authority of the state, and steadily less responsive to internal organizational complexity. We expect to find increasing correlations between the Childhood Index and measures of state authority, but decreasing correlations between this index and measures of social and economic development.

RESULTS I:

CHANGES IN THE CONSTITUTIONAL DEFINITION OF CHILDHOOD, 1870-1970

The American reader lives in a society with a federal state whose formal authority is somewhat restricted, though this authority has steadily expanded through interpretive as well as formal changes. Unlike virtually all other constitutions currently in force, the United States Constitution contains no provisions distinguishing children as a separate category of persons with unique rights or obligations, or subject to particular state action. In American society childhood is for the most part officially under the jurisdiction not of the federal state but of various political subunits, a very unusual situation. Hence the USA receives scores of zero on all of the indicators discussed below, including the Childhood Index.

To explain the coding scheme used to study the constitutional definition of childhood, we present in the Appendix all of the references to children and the state's functions vis-à-vis children from a constitution that rather fully expresses childhood ideology. We refer to the relevant passages, which are lettered consecutively in the left margin, in presenting the coding system. The particular constitution we have chosen, that of Ecuador of 1945 (in force until 1972), is typical of constitutions that clearly embody the ideology of childhood we have discussed above.

The nine distinct indicators used for coding constitutions are as follows:

A. Childhood as a distinct state of biography. Does the constitution define childhood as a distinct category apart from other stages in life? If so, how explicitly does it do so? This four-point scale includes:

- (0) There is no mention of childhood or children;
- (1) One or more general references to childhood are made;
- (2) Reference is made to the category and explicit age limits are stated (e.g., "under 18");
- (3) Age limits are explicitly stated for *two or more* segments of the childhood category (e.g., "under 14" and "14 to 18").

In the Appendix, note that the first reference to children in Ecuador's constitution occurs in paragraph (A), but that later in paragraphs (P) through (R), more explicit references appear. Children under 12 are clearly differentiated from children between 12 and 14 or those between 14 and 18, and the state declares that these separate segments within the general category shall be treated differently in society. We code this constitution (3), then, on the first indicator. A score of (2) would result if only one age limit for the category were stated, as in the 1960 constitution of Cyprus in which reference is made only to children before they "attain the age of sixteen" (Peaslee, 1965-70).

B. State responsibility for children. Does the state have responsibility for the child, or is such responsibility left with the family? This indicator is coded as follows:

- (0) No mention is made of state responsibility for children;
- (1) There is a general statement of state protection for children;
- (2) Ultimate responsibility for children rests explicitly with the state (in case of loss of parents or extreme poverty, for example);
- (3) Specific protective agencies or rules are described that have the task of putting point (2) into effect; orphanages, child care centers, etc.

Paragraph (B) of the Appendix shows immediately that the state must offer general protection for children, a score of (1). But paragraph (C) goes further, showing the conditions under which the state as-

sumes responsibility for children. This example obtains a score of (2) because no specific agencies are identified that will satisfy point (3). The 1946 constitution of Albania takes the further step in the following passage (Peaslee, 1965-70):

Article 17. The State gives special protection to the interests of mother and child by . . . setting up homes for expectant mothers and homes for bringing up and sheltering children.

We score this passage (3). As a final example, a clause that shows up frequently in Catholic countries receives a score of (1):

The protection of youth against exploitation and against moral, intellectual, and physical abandonment, is an obligation of the State and of the public bodies. (1961 constitution of Gabon; Peaslee, 1965-70)

Where childhood exists as a distinct category in constitutions and the state has some measure of responsibility for children, two institutional areas primarily reflect this jurisdiction: labor legislation and education. In these areas, the state has two responsibilities: to protect children, who are considered weak, innocent, tender, or helpless; and to shape children in certain collectively-desired ways. This is possible according to socialization ideology because children are clearly unformed "personalities" at birth whose "aptitudes" require "development" through the process of "growth" (cf. paragraphs C, G in the Appendix). Labor legislation serves to keep children out of the job market and in the school, where state-directed socialization takes place. The state can there ensure that children become "socially useful," with "a democratic spirit of Ecuadorianism and of human solidarity," that is, supportive of the nation and the state, as well as of broader "human" values (cf. paragraphs F, J). State functions are coded as follows:

C. Child labor regulation.

- (0) No mention is made of the function;
- (1) The state has general authority to regulate labor of all types (which would include child labor);
- (2) The state has explicit authority to regulate child labor in particular;
- (3) Specific rules regulating child labor are given.

Article 148 in the Appendix gives a long list of state regulatory powers over labor, with paragraphs (P) through (R) the key for this variable. Ecuador's constitution obtains a score of (3) because there are basic rules regulating child labor as part of the constitution (not just as part of labor legislation, which is more easily changed). A weaker statement appears in the 1937 constitution of Ireland, which obtains a (2):

Article 45 (4) 2. The State shall endeavor to ensure that the strength and health of workers, men and women, and the tender age of children shall not be abused and that citizens shall not be forced by economic necessity to avocations unsuited to their sex, age, and strength. (Peaslee, 1965-70)

Finally, a series of indicators assesses the degree of state control of education and the degree to which children must participate in the national educational system. These are straightforward.

D. State control of education. How strong is the authority over education given to the state?

- (0) No authority over education is mentioned;
- (1) The state has authority to regulate or directly operate schools, but nonstate schools also are permitted;
- (2) The state has monopoly power over education: nonstate schools are prohibited.

In the Appendix, private education is under state regulation but is permitted (paragraphs E, F), although it cannot receive support from the state (paragraph I). A score of (1) is assigned. An example of a stronger statement comes from the 1954 constitution of China (Peaslee, 1965-70):

Article 94. Citizens of the People's Republic of China have the right to education. To guarantee enjoyment of this right, the state establishes and gradually extends the various types of schools and other cultural and educational institutions.

Private schools are not explicitly outlawed here, but it is clear from the context that only the state may operate schools.

E. Obligation of the state to provide education. Whether or not the state has authority over education, does it also have the obligation to provide it? Almost all constitutions agree that it does. Scores include:

- (0) No obligation is mentioned;
- (1) The state has a general obligation either to regulate or to operate schools, but the scope of the obligation is less than point (2);
- (2) The state's obligation extends to all children; usually, it must guarantee education to all children.

Paragraph (B) most clearly identifies the state's obligation for education in the Appendix: the "rights . . . to education . . . are guaranteed." In conjunction with the first line of Article 143, we code this example (2).

F. Levels of education specified. We code:

- (0) Education is not mentioned;
- (1) Education is mentioned, but no levels are specified or primary education alone is specified;
- (2) Two or more levels are specified (primary, secondary, or tertiary), or primary is assumed by the context and secondary or tertiary is specified.

This indicator, simply describing the extent of the schooling system, shows how much of the child's early socialization is explicitly within the state's jurisdiction. The example is coded (2), because both primary and tertiary levels receive mention (paragraphs I, K).

G. The right to education. Rights, in the modern polity, have two facets: they represent claims that individuals can make on the state, and they imply an obligation of the state to satisfy those claims (a point to which we return). What claims can child-citizens make on the state with regard to education? We code:

- (0) No right to education is mentioned;
- (1) A right is stated in general terms;
- (2) A right to a specific level or period of years is stated.

Note that it was not necessary to make provision for cases in which the right to education applies to only part of the population, for it is everywhere stated as a completely universal right if it exists at all. In the example, paragraph (B), defining this right, receives a score of (1). An example that defines a more explicit right is that of Denmark (constitution of 1953; Blaustein and Flanz, 1971-): "Article 76. All children of school age shall be en-

titled to free instruction in the elementary schools. . . ."

H. The duty to be educated. In almost every constitution, the right to education is accompanied by a duty to be educated as well. Coding parallels the previous indicator:

- (0) No duty to be educated is mentioned;
- (1) A duty is stated in general terms;
- (2) A duty to complete a specific level or period of years is stated.

The Ecuadorian constitution specifies obligatory elementary education, a code of (2), in paragraph (I). A weaker statement appears in the Taiwan constitution of 1946, in which right and duty are inseparable. This is a typical passage (Blaustein and Flanz, 1971-): "Article 2. The people shall have the right and the duty of receiving citizens' education."

Table 1 presents summary results of the application of this scheme to the constitutions of all independent countries at 20-year intervals between 1870 and 1970. For each indicator we present in the upper row the mean for all countries at each time. Thus, for the first indicator, childhood as a distinct stage of biography, the upper row shows that the 43 nations of 1870 had a mean of .42 (where the maximum value is 3), while by 1970 the mean for the 139 independent nations existing at that time had risen to 1.2. The increase occurred mainly between 1930 and 1950, a period of almost universal revision of constitutions as well as considerable state formation in previously colonized areas. Other indicators, such as the duty to be educated (H), increase more uniformly. Note that after 1950 there is rather little increase in most indicators, and some even decline slightly.

The lower row for each indicator shows the mean for a fixed panel of countries. The panel consists of all those countries existing in 1870 that remained independent through 1970 (or were subject to only temporary occupation, such as Japan from 1946-55). Thus, for example, only one of these nations' constitutions contained any provisions for state responsibility for children (mean of .02) in the years 1870 to 1910. By 1930 such provisions had begun to appear more frequently in constitutions and by 1970 the mean had risen to .89,

Table 1. Mean Scores on Indicators of Constitutional Differentiation of Childhood, 1870-1970 (for All Countries in Existence at Any Given Time, and Separately for Countries Independent by 1870)

Indicator		Time					
		1870	1890	1910	1930	1950	1970
a. Childhood as a distinct stage of biography (scored 0-3)	—all cases	.42	.30	.36	.49	1.0	1.2
	—independent by 1870 only	.42	.33	.40	.48	1.0	1.0
b. State responsibility for children (scored 0-3)	—all cases	.02	.02	.04	.19	.88	1.2
	—independent by 1870 only	.02	.02	.02	.21	.93	.89
c. Child labor regulation (scored 0-3)	—all cases	.09	.15	.15	.59	1.2	1.0
	—independent by 1870 only	.09	.17	.16	.61	1.3	1.4
d. State control of education (scored 0-2)	—all cases	.72	.78	.79	.89	1.0	1.0
	—independent by 1870 only	.72	.81	.81	.86	1.0	1.0
e. Obligation of the state to provide education (scored 0-2)	—all cases	.61	.72	.68	.76	1.3	1.0
	—independent by 1870 only	.61	.76	.77	.77	1.4	1.1
f. Levels of education specified (scored 0-2)	—all cases	.95	.96	1.0	1.2	1.4	1.2
	—independent by 1870 only	.95	1.0	1.0	1.1	1.5	1.5
g. Right to education (scored 0-2)	—all cases	.54	.85	.83	.85	1.3	1.2
	—independent by 1870 only	.54	.83	.88	.73	1.3	1.3
h. Duty to be educated (scored 0-2)	—all cases	.23	.61	.70	.91	1.2	1.1
	—independent by 1870 only	.23	.57	.63	.93	1.3	1.3
N for all cases		43	46	53	65	77	139
N for countries independent by 1870 only		43	43	43	43	43	43

with many constitutions now making such provisions. In general, there is very close correspondence between the pattern followed by this fixed panel, in the lower rows, and that of the entire set of nations, in the upper rows. This indicates that the increases in the indicators are not confined only to the older, more established nations, a point elaborated below. Overall, Table 1 provides initial evidence in support of our first hypothesis: The ideology of childhood has been incorporated in the structure of the state throughout the world.

Index of Constitutional Provisions Concerning Childhood (The Childhood Index)

To summarize the individual indicators,

we have combined a subset of them as an overall index of the degree of differentiation of childhood in relation to the state. The Index includes the following:

- Childhood as a distinct stage (indicator A above);
- State responsibility for children (B);
- Child labor regulation (C);
- Obligation of the state to provide education (E);
- Levels of education specified (F);
- The duty to be educated (H).

The maximum score on the Childhood Index is 15; the minimum is zero. Indicators D and G, state control of education and the right to education, are omitted because they are largely redundant with indicators E and H; in practice, our results would look no different if we had included D and G rather than E and H.

To justify combining the individual indicators into the more general Childhood Index, we examined item intercorrelations. The indicators are indeed positively intercorrelated, and significantly so. Over the period of our study, the mean item intercorrelation is .33. Of the 90 item intercorrelations included, only seven are zero or negative, and no pair of items has a consistently negative intercorrelation over the six times. Sixty-three of the intercorrelations take a value of .2 or greater.

More significantly, these data show that the state-organized definition of childhood becomes an increasingly coherent ideology over time. The mean item intercorrelations for 1870 and 1890 are only .26 (43 and 46 cases). By 1950 and 1970, the mean intercorrelations had risen to .52 and .89 (77 and 139 cases).⁶ The inclusion of constitutional provisions for childhood decreasingly becomes, we infer, a matter of accidental choice, and increasingly reflects a coherent world view.

Table 2 sums up the detailed results of Table 1. We present, in the upper section, the mean scores for all countries at each time on the Childhood Index. The overall increase in the Index is dramatic—it approximately triples over the century, with most of the increase occurring between 1910 and 1950. This result conforms to the first hypothesis as well.

Further support is provided by the middle section of Table 2. Here we present country scores for constant panels of countries distinguished according to date of independence. Each row of this panel shows actual *changes* in constitutional scores on the Index for each group of countries.⁷ Constitutional changes are more common in the world than an American reader might anticipate, and the

data reflect this. Countries almost continuously independent since 1870 show very large, statistically significant, and rather consistent increases over time. So do the countries becoming independent between 1870 and 1910 and those becoming independent between 1910 and 1930. Only the group becoming independent between 1930 and 1950 shows no overall increase on the index. These data clearly establish that the increases in national scores on the Index of Constitutional Provisions Concerning Childhood do *not* simply reflect the addition of new countries to the population of independent nations: giver groups of countries change in their scores over time.

The data reveal a further important result. Each group of new nations entering the world system of nation-states enters with a higher score on the Childhood Index than did preceding groups. National constitutions do not simply reflect processes of internal development. They reflect legitimating ideas dominant in the world system at the time of their creation (see Stinchcombe, 1965, for an organizational analogue). Each new group of countries formulates a set of constitutions almost indistinguishable, in provisions concerning childhood, from the constitutions evolving in groups of older countries. This finding suggests strongly that the expansion of state authority over childhood is not a simple reflection of the level of social organization in each country. More recently independent countries tend to be less developed, less internally differentiated, and more peripheral in the world system. Yet their formal state structures reflect the same claims for responsibility for childhood that characterize older, more central nations.

The bottom panel of Table 2 approaches this issue in another way. Countries are distinguished according to their centrality in the world system, defined in terms of the proportion of world trade in which they engage (Banks, 1975). Countries with over 3% of world trade are defined as central, and those with less than 1% are defined as peripheral. This measure is strongly related to other measures of power and centrality in the world system (cf. Boli-Fennett, 1977).

⁶ Item intercorrelations also rise over time when we examine data for the constant set of countries in existence since 1870.

⁷ Countries comprising each panel are identified in Boli-Bennett (1976). In general, the oldest panel consists mainly of western European and Latin American countries; the 1871–1910 panel is a highly disparate group; the 1911–30 panel is largely eastern European countries; the 1931–50 panel, Middle Eastern and Asian countries; and the 1951–70 panel, mostly African and a few Asian countries.

Table 2. Mean Scores on Index of Constitutional Provisions Concerning Childhood (Cases in Parentheses)

	1870	1890	1910	1930	1950	1970
(A) Mean score for all countries independent at each time	2.33 (43)	2.76 (46)	2.93 (53)	4.11 (65)	7.03 (77)	6.78 (139)
(B) Mean score by date of independence:						
Countries independent by 1870 (43 cases)	2.38	2.86	3.00	4.11	7.54	7.21
Countries independent 1871-1910* (8 cases)			2.63	3.63	7.38	8.25
Countries independent 1911-1930 (10 cases)				4.44	5.50	8.10
Countries independent 1931-1950 (14 cases)					6.00	5.69
Countries independent 1951-1970 (57 cases)						6.28
(C) Mean scores by centrality in world system:						
Countries controlling 3% or more of world trade (central)	1.38 (8)	2.00 (8)	2.00 (9)	3.75 (8)	4.38 (8)	5.20 (10)
Countries controlling 1-3% of world trade (semiperipheral)	2.33 (6)	2.89 (9)	2.27 (11)	4.07 (15)	7.81 (16)	6.56 (9)
Countries controlling less than 1% of world trade (peripheral)	2.45 (29)	2.93 (29)	3.39 (33)	4.19 (42)	7.19 (53)	6.93 (120)

* Only 2 countries became independent between 1870-1890. They are combined with the 6 countries independent between 1891-1910.

Notes on statistical significance of difference of means:

1. All changes between the first and last date for every row in the table are significant (t-test, $p < .05$; for most changes, $p < .01$) except for the decrease in the mean of countries independent 1931-50.
2. One-way analysis of variance shows significant differences ($p < .01$) among initial scores of panels of countries in (b), though not all pairs of initial scores differ significantly from each other.
3. Analysis of variance shows no significant cross-sectional differences among panels of countries in (b) and no significant differences among pairs of means in (b), for any of the six coding times.
4. Analysis of variance shows weakly significant difference ($p < .10$) among sectors of the world system in (c) in 1910 and 1950; at other times these differences are not significant.
5. These scores represent not samples of a population but the population itself. Sampling error is eliminated, though some coding error persists. The significant tests presented here apply more stringent standards for inferring differences among means than may be necessary, so that rather small differences probably are less likely to be due to error than the tests suggest.

All three categories of countries show the increases in mean scores on the Childhood Index that we have come to expect. Strikingly, however, the most peripheral countries tend to show the highest, not the lowest, scores. The expansion of state authority over childhood is not simply a matter of the diffusion of rules from centers to peripheries, or of copying of the constitutions of metropolitan powers. This expansion is, rather, a matter of world ideology

that peripheral countries adopt most completely. Central countries are a little slower to adopt it. Perhaps this is because they retain more stability and autonomy in their political structures, but it also may reflect the fact that the differentiation of childhood arose from the social organization of core countries and is less a special purposive project of central states than of the mobilizing peripheral states. The social ideas differentiating childhood from

adult life are not fully incorporated in central states. Peripheral countries institutionalize these ideas as ideological rules somewhat more directly.

The results of Table 2 provide strong support for our first hypothesis. Constitutional rules concerning childhood expand very rapidly in the present period and penetrate the state systems of most societies, both new and old, central and peripheral.

RESULTS II: CORRELATIONS WITH STATE POWER AND SOCIETAL COMPLEXITY

Our remaining three hypotheses embody the following ideas: Constitutional rules incorporating the ideology of childhood in the state structure reflect the general authority of the state, and they increasingly reflect this authority over time. They reflect much less, and reflect decreasingly over time, the complexity of the social organizational systems of societies, both their level of technical development and the organizational power of the state (as distinct from its authority). Correlations between the Childhood Index and the indicators described below provide tests for these ideas.

(A) General State Authority

We focus here on state authority, not the concrete organizational power of the state bureaucracy. The organizational power of the state may vary greatly from its ideologically legitimated authority for many reasons—most obviously, economic development and centrality in the world system tend to build organizational power into the state structure without necessarily increasing state authority. Two measures are relevant:

Index of constitutionally defined state authority. This index derives from a very comprehensive coding of constitutions for the areas of jurisdiction they explicitly assign to the state. Most constitutions contain "legislative lists" of the institutional areas over which the state may exercise control, and the chief purpose of constitutions is to spell out the relationship between state and society implicit in these lists. The index simply sums the number of institutional areas under state jurisdic-

tion and the number of administrative mechanisms available to the state to manage society. Examples of the institutional areas include agriculture, drugs and alcohol use, foreign trade, education, health, labor disputes, mining and the professions (a total of 56 items). Examples of administrative mechanisms include press censorship, civil service, emergency powers, immigration, patents, nationalization of industry, police, public works, welfare programs, and the like (45 items). For a fuller discussion see Boli-Bennett (1976).

Index of constitutionally defined citizen rights. Far from being abstract properties inhering in individuals, constitutionally defined rights of citizenship specify the claims citizens can make on the state. They have effective meaning only when the state is obligated to protect or fulfill them (Marshall, 1948; Duchacek, 1973). Thus, they provide a separate indicator of state authority, this one more from the "bottom up" than from the "top down." This index sums all citizen rights enumerated in each constitution, usually an easy task because most constitutions contain a separate section on citizen rights. Examples of common rights include the rights to education, health care, leisure, voting, social security, and equal wages for equal work, as well as the more familiar rights of assembly, association, due process, and property (a total of 38 items).

(B) The Organizational Complexity of Society

Here we are interested in measures of technical development, or what is often called *social and economic development* or *modernization*. Limited by the availability of data, we employ the following mostly conventional measures:

Urbanization: the proportion of the population living in places of 20,000 or more inhabitants, taken from Banks (1975).

Nonagricultural labor force as a proportion of total labor force: the extent to which the economy emphasizes secondary and tertiary activity. Data are taken from Bairoch (1968), Clark (1957), Banks (1975), Howard (1935), and Feuerwerker (1968).

National income per capita: an approximation of total economic exchange activity, and thus a reflection of the degree of division of

labor, from Clark (1957), Banks (1975), and Feuerwerker (1968; 1969).

Technical development index: a variable constructed from a factor analysis of a number of measures of industrialization and complexity. The index is the factor score for each case calculated by using the loadings on a single factor that combines per capita energy consumption, the percentage of children enrolled in primary schools, and the three indicators given above. Details are available in Boli-Bennett (1976).

The intercorrelations among these indicators are all high, generally ranging between .50 and .80 over most of the period. The technical development index correlates most highly with urbanization but all variables that comprise it have very high loadings on the factor.

(C) State Organizational Power

The state has effective organizational control of society to the extent to which it can extract resources from society and use them for state programs. The best available measure of this concept also comes from Banks (1975): *government*

revenue as a proportion of national income.

Table 3 shows the correlations of each of the seven variables discussed above with the Childhood Index, for each of the six times from 1870 to 1970.⁸ Consider first the data, in the first panel of Table 3, on the two measures of general state authority. The hypotheses are obviously confirmed: The constitutional elaboration of rules distinguishing childhood is highly correlated with, and increasingly correlated over time with, these two measures. As the ideological legitimization of the state expands, so do rules distinguishing child-

⁸ It could be argued that we should employ raw regression slopes in these analyses to facilitate the comparability of data over time. We think this is inappropriate here—means, ranges and variances of the independent variables change so much, and change in social meaning so much, as to make comparisons in regression slopes difficult to interpret. Correlations artificially standardize for these changes in meaning treating variances as if they were conceptually constant over the period. This is, in our view, a sensible first approximation. The general effects discussed here, however, would remain if regression slopes were employed.

Table 3. Correlations between Index of Constitutional Provisions Concerning Childhood and (A) Measures of General State Authority, (B) Measures of Social Organizational Development or Complexity, and (C) State Organizational Power (Cell Entries are Person Zero-Order Correlations; Cases in Parentheses)

	1870	1890	1910	1930	1950	1970
A. General state authority:						
—Index of constitutionally defined state authority	.49** (38)	.55** (42)	.58** (49)	.72** (61)	.77** (76)	.72** (139)
—Index of constitutionally defined citizen rights	.48** (38)	.54** (42)	.59** (49)	.60** (61)	.76** (76)	.60** (139)
B. Social organizational complexity or development:						
—Urbanization	.21 (41)	.10 (44)	-.06 (48)	.00 (61)	-.04 (75)	.10 (135)
—Nonagricultural labor force as percent of total labor force	.44 (12)	.61* (15)	.20 (24)	.09 (38)	.33** (63)	.20 (65)
—National income/capita	.42 (16)	.23 (19)	.29 (23)	-.27 (30)	-.24 (62)	-.20* (113)
—Technical development factor	-.14 (40)	-.34* (42)	-.38* (45)	.04 (49)	-.08 (70)	.08 (134)
C. State organizational power:						
—Government revenue/national income	-.06 (16)	-.05 (19)	-.26 (23)	-.10 (26)	-.10 (56)	-.07 (112)

* $p < .05$

** $p < .01$

dren as a status group under state jurisdiction.⁹

The second panel of Table 3 shows a very different picture. The four different measures of societal complexity generally show low (often negative) correlations with the constitutional differentiation of childhood. Further, these correlations decline over time and almost none of them are significant. The average correlation declines from .21 and .15 at the first two times to -.01 and .04 at the last two times. There is some evidence, in other words, that aspects of technical development, particularly national wealth and urbanization, may be initially responsible for the elaboration of ideological rules in national constitutions distinguishing children; but more recently, political ideologies defining the state as responsible for childhood have become less tied to social organizational constraints. These ideologies are worldwide creatures most fully expressed in societies (often peripheral ones) where the general depiction of the state as ultimately responsible for social life and social progress is strongest.

The final panel of Table 3 shows that the Childhood Index is virtually unrelated to state organizational power, as measured by government revenue as a proportion of national income, over the entire period. That is, whether or not the state has effective ability to implement the organizational forms that would bring the ideology to life, its authority over childhood may be strongly institutionalized. This development helps pave the way for the organizational forms that the state eventually establishes.

Our last three hypotheses, then, receive substantial confirmation. Constitutional rules differentiating childhood are indeed strongest (and increasingly so) in societies with highly expanded state authority. They are much less, and decreasingly, affected by social organizational complexity or technical development. They also are largely unaffected by state organizational power.

⁹ These results, like all of those that follow, are reproduced almost exactly in analyses using data concerning only the panel of countries existing continuously since 1870. The changing correlations thus are not due simply to changing sets of countries.

Comparisons among types of countries. The institutional rules incorporating childhood in the state structure have spread rapidly over the last century and they are closely associated with expanding general state authority. Their presence is not much related to the complexity or development of the social organizational structure of society, but there is much variation among countries to be explained. To begin to explain this variation, we have examined the 1970 country scores on the Childhood Index. Table 4 reports means for groups of countries, along with means on the Index of Constitutionally Defined State Authority. The differences among the types of countries identified are large and statistically significant. What do they reflect?

Two factors associated with differences in abstract political ideology clearly help explain variations among countries. (a) In Communist countries, which are organized around an ideology emphasizing very aggressive state action in society, we find very high scores on the Childhood Index, and also on the Index of State Authority. (b) In Catholic countries, state au-

Table 4. Means, Index of Constitutional Provisions Concerning Childhood and Index of Constitutionally Defined State Authority, 1970, for Major Types of Countries

Type of Country	Childhood Index	State Authority Index	Cases
Latin American	9.9 ¹	63.9 ¹	19
Communist-party dominant	9.4	69.3	14
European Catholic	7.9	44.4	9
African, Catholic colonizer	6.6	34.6	18
Moslem or Middle Eastern	6.5	43.2	21
Asian	6.0	48.9	16
West Indies	5.3	37.8	4
African, Protestant colonizer	5.0	38.2	16
European Protestant	4.2	48.2	17

Note: *European Protestant* includes technically developed countries originally colonized by Britain: USA, Australia, etc.

¹ One-way analysis of variance shows significant differences ($p < .01$) among types of countries in both the Childhood Index and the State Authority Index.

thority derives partly from a natural law tradition and the wide authority of corporate groups within society (Swanson, 1967). European Catholic countries have higher scores on the Childhood Index than European Protestant ones. African countries originally colonized by European Catholic powers (France and Belgium) similarly have higher scores than African countries colonized by European Protestant powers (Great Britain). This difference might be larger except for the fact that both France and Belgium have the lowest scores among European Catholic countries (four and one, respectively), and low scores are reflected in their colonies. Latin American countries, on the other hand, were colonized by Spain and Portugal, which have high scores (eleven and ten), and their scores are correspondingly higher.

The constitutions of former colonies, however, do not simply reflect those of their metropolitan powers. Former French colonies have much higher scores than does France, and Britain's constitution provides a poor model to copy because it is uncoded (an extreme rarity in the contemporary world). Newer countries often have higher scores than the older core powers, which retain constitutional structures from the more distant past. The diffusion of political ideology is much more complex than the simple borrowing of constitutional provisions. Ideologies built into the social organizational life of core powers tend to become constitutional principles in peripheral nations, often well before they are incorporated fully in the constitutions of the core powers themselves.

In general, differences on the Childhood Index are paralleled by differences on the Index of State Authority. Note, however, that Catholic countries and their ex-colonies, while higher on the Childhood Index, are generally lower than their Protestant counterparts in general state authority, with the exception of Latin American countries.

CONCLUSION

We have measured the elaboration of constitutional rules defining the nation-

state as responsible for the distinct category of persons called *children* in the contemporary world system. Childhood increasingly comes under the jurisdiction of the state, a process that parallels the generally increasing power and authority of the state. The institutional rules of state-managed childhood spread through the world system, not depending much on particular problems or organizational features found in specific national societies. Rather, they diffuse to all types of states as an element of prevailing political ideology.

The dominance of the ideology of differentiated and state-managed childhood reflects the rise of both individualism and the rationalized nation-state. The individual's status in society increasingly takes the form of membership (citizenship) in the nation and state: the individual becomes as much an agent of the collectivity as of personal or subgroup interests. In this vein, modern childhood comes under the authority of the state as part of the state's command over the socialization of its constitutive members—its authority to prepare its citizens for their roles in aiding national development, achieving progress, and obtaining success in the world system.

The diffusion of this ideology clearly has been aided by the integration and, in some key respects, homogeneity of national societies in the globally unified value and stratification system. States pursue progress within the agreed-upon frame. Among these formally similar states such institutions of modernity as collectively managed childhood, evolved organizationally in world centers and moved rapidly as institutional recipes for progress to the most distant peripheries.

We study here ideology, but worldwide ideologies penetrate and structure local organizational life. States, not only in theory, but also in practice, extend their authority throughout the social structure; they build the concrete organizations which manage the lives of children. The process we depict here parallels—and we believe it tends to create—the massive extensions of collective authority over children found in such phenomena as universal schooling, which has rapidly become a

worldwide fact and is decreasingly related to variations in national economic and social organization (Meyer et al., 1977). Ideology paves the way for the application of state programs and the universalization of state-directed socialization. We should thus expect further growth of other organizational forms participating in this process: national child health care, juvenile courts, state-employed child psychologists, and so on.

Whether this process is likely to "liberate" children or is instead a repressive disaster is hard to say. As with many other features of technical society, we witness here the substitution of a rationalized locus of control (the state) for a more traditional locus (the family). In our view it seems unlikely that children will better control their own lives if, for example, we eventually institute one of the cherished aims of the children's rights movement: full political participation, including the right to vote, for children from the moment of birth (cf. the section on "bills of rights" in Gross and Gross, 1977). The extension of citizenship, to the kindergarten or anywhere else, is at best an ambiguous step on the path to true equality or freedom.

APPENDIX

REFERENCES TO CHILDHOOD AND STATE RESPONSIBILITIES FOR CHILDREN IN THE 1945 CONSTITUTION OF ECUADOR

The Family

Article 142. The State protects the family, marriage, and motherhood. . . .

(A) Illegitimate children have the same rights as legitimate children in terms of upbringing, education, and inheritance. . . .

(B) The defense of the physical, mental and moral health of infancy and the rights of the child to education and a home life are guaranteed.

(C) The State shall create for minors who lack family or economic protection, conditions adequate for their growth.

(D) In penal matters, minors shall be subject to special legislation, which shall be protective and nonpunitive.

Education and Culture

Article 143. Education is a function of the State.

(E) Private education is guaranteed; it must be adapted to the laws and to the regulations and official programs.

(F) Public and private education have as their ob-

ject to make the students socially useful. They must be imbued with a democratic spirit of Ecuadorianism and of human solidarity.

(G) Public education . . . shall be organized in such a way that there is adequate expression and continuity in all grades. It shall employ methods that are based on the activity of the student and that develop his aptitudes while respecting his personality.

(H) Official education is secular and gratuitous in all its grades. Neither the State nor the Municipalities can subsidize any other educational system; but the social services shall be supplied without discrimination to all students who need them.

(I) Elementary education is obligatory. The State shall provide the school materials necessary in public education, without cost to the student. . . .

(J) The State shall devote special attention to development of technical education, in accordance with agricultural and industrial needs.

(K) The universities are autonomous, according to the law, and shall devote special attention to the study and resolution of national problems and to the dissemination of culture throughout the popular classes. To guarantee that autonomy, the State shall support the creation of a university endowment. . . .

(L) The State shall aid students in need in order to facilitate their complete education.

(M) An item in the budget shall be allocated annually for fellowships for children of workers, artisans, and peasants.

(N) The law shall determine the manner in which students shall intervene in directive and administrative affairs of institutes of education.

Labor and Social Security

Article 148. The basic standards that regulate work in Ecuador are as follows:

(O) f. The State shall seek to establish the family wage, preferably using the system of child subsidy;

(P) i. The maximum work period shall be eight hours. . . . The night shift shall be shorter in length than the day shift and remunerated with more salary and neither women nor minors of less than eighteen years of age may be employed on it. . . .

(Q) o. Labor of persons under fourteen years of age is prohibited save for the exceptions established by law, and labor by persons under eighteen years of age shall be regulated.

(R) p. It is prohibited to hire minors up to twelve years of age as domestic servants. . . .

Source: Blaustein and Flanz (1971-)

REFERENCES

- Ariès, Philippe
1962 *Centuries of Childhood*. New York: Vintage.
- Arosemena, Justo
1870 *Constituciones Políticas de la América Meridional*. Le Havre: Lemale.
- 1888 *Estudios Constitucionales sobre los*

- Gobiernos de la América Latina. 2 Vols. Paris: Roger y Chernoviz.
- Bairoch, P.
1968 *La Population Active et Sa Structure*. Bruxelles: Université Libre de Bruxelles.
- Banks, Arthur S.
1975 *Cross-National Time-Series Data Archive*. Binghamton: Center of Comparative Political Research, State University of New York.
- Bendix, Reinhard
1964 *Nation-Building and Citizenship*. New York: Wiley.
- Berger, Peter, Brigitte Berger, and Hansfried Kellner
1973 *The Homeless Mind*. New York: Random House.
- Blaustein, Albert P., and Gisbert H. Flanz
1971- *Constitutions of the Countries of the World*. 14 Vols. Dobbs Ferry: Oceana Publications.
- Boli-Bennett, John
1976 *The Expansion of Nation-States, 1870-1970*. Ph.D. dissertation, Department of Sociology, Stanford University.
1977 "Global integration and the rise of the universal omnivorous state." Paper presented at the annual meeting of the Pacific Sociological Association, Sacramento.
- Clark, Colin
1957 *The Conditions of Economic Progress*. 3rd ed. London: Macmillan.
- Dareste, F. R., with P. Dareste.
1891 *Les Constitutions Modernes*. 2 Vols. 2nd ed. Paris: Challamel.
1910 *Les Constitutions Modernes*. 2 Vols. 3rd ed. Paris: Challamel.
1928- *Les Constitutions Modernes*. 6 Vols. 4th ed. Ed. by Joseph Delpech and J. Laferrière. Paris: Sirez.
- de Mause, Lloyd
1974 *The History of Childhood*. New York: Psychohistory Press.
- Duchacek, Ivo K.
1973 *Rights and Liberties in the World Today*. Santa Barbara: ABC-CLIO.
- Ellul, Jacques
1967 *Métamorphose du Bourgeois*. Paris: Calmann-Levy.
1975 *Trahison de l'Occident*. Paris: Calmann-Levy.
- Feuerwerker, Albert
1968 "The Chinese economy, 1912-1949." Ann Arbor: Center for Chinese Studies, University of Michigan.
1969 "The Chinese Economy, ca. 1870-1911." Ann Arbor: Center for Chinese Studies, University of Michigan.
- Gross, Beatrice, and Ronald Gross
1977 *The Children's Rights Movement*. Garden City: Anchor.
- Holt, John
1973 *Escape from Childhood*. New York: Simon and Schuster.
- Howard, Louise
1935 *Labor in Agriculture: An International Survey*. London: Oxford University Press.
- Inkeles, Alex, and David Smith
1974 *Becoming Modern*. Cambridge, Ma.: Harvard University Press.
- Marshall, T. H.
1948 *Citizenship and Social Class*. New York: Doubleday.
- Meyer, John W., John Boli-Bennett, Christopher Chase-Dunn
1975 "Convergence and divergence in development." *Annual Review of Sociology* 1: 223-46.
- Meyer, John W., and Joane P. Nagel
1975 "The changing status of childhood." Paper presented at the annual meeting of the Society for the Study of Social Problems, San Francisco.
- Meyer, John W., Francisco O. Ramirez, Richard Rubinson, and John Boli-Bennett
1977 "The world educational revolution, 1950-1970." *Sociology of Education* 50:242-58.
- Peaslee, Amos J.
1950 *Constitutions of Nations*. 3 Vols. Concord: Rumford Press.
1956 *Constitutions of Nations*. 3 Vols. 2nd ed. The Hague: Martinus Nijhoff.
1965- *Constitutions of Nations*. 4 Vols. Rev. 3rd ed. Ed. by Dorothy Peaslee Xydis. The Hague: Martinus Nijhoff.
- Posener, Paul
1909 *Die Staatsverfassungen des Erdballs*. Charlottenburg: Fichtner.
- Ramirez, Francisco O., and Richard Rubinson
1979 "Creating members: the national incorporation of education." Chap. 4 in John W. Meyer and Michael T. Hannan (eds.), *National Development and the World System: Educational, Economic, and Political Change, 1950-1970*. Chicago: University of Chicago Press. In press.
- Skolnick, Arlene
1975 "The limits of childhood: conceptions of child development and social context." *Law and Contemporary Problems* 39:38-77.
- Stinchcombe, Arthur L.
1965 "Social structure and organizations." Pp. 153-93 in James G. March (ed.), *Handbook of Organizations*. Chicago: Rand McNally.
- Swanson, Guy E.
1967 *Religion and Regime*. Ann Arbor: University of Michigan Press.
1971 "An organizational analysis of collectivities." *American Sociological Review* 36:607-24.
- Tyack, David
1966 "Forming the national character." *Harvard Educational Review* 36:29-41.
1974 *The One Best System*. Cambridge, Ma.: Harvard University Press.
- Weber, Max
1930 *The Protestant Ethic and the Spirit of Capitalism*. New York: Scribner's.

BRINGING THE BOSS BACK IN: EMPLOYER SIZE, EMPLOYEE SCHOOLING, AND SOCIOECONOMIC ACHIEVEMENT*

ROSS M. STOLZENBERG

University of Illinois

American Sociological Review 1978, Vol. 43 (December):813-828

Employers play a crucial role in the process of social stratification in the U.S.: the jobs they provide are the primary mechanism by which individuals are distributed among occupations and by which earnings are distributed among persons. But the vast majority of sociological work on socioeconomic achievement ignores employers completely, and nearly all related work by economists either ignores employers or else ignores characteristics of workers, substituting one omission for another. In the theoretical section of this paper, we review and combine sociological work on organizational structure, sociological studies of social stratification, and economic research on labor markets, industrial wage differentials and human capital. In this review, we hypothesize that the size of an employer organization works indirectly through other dimensions of organizational structure to alter the effect of workers' schooling on their wages and occupational attainment. In particular, we hypothesize that the effect of workers' schooling on earnings and the effect of workers' schooling on occupational SES increase as logarithmic functions of the size of the organization which employs them. Since the relationship between schooling and occupational attainment is a central part of social stratification research, and because the effect of schooling on earnings is a fundamental feature of sociological and human capital research on earnings, this hypothesis links organizational structure to the very heart of current issues in social stratification and human capital studies. In the empirical section of the paper, we divide a national probability sample of workers into five groups, depending on the size of the establishment in which they work. Fitting models of earnings and occupational attainment in each of these groups and then relating schooling effects to the size of the establishment which defines the groups, we find that the effect of workers' schooling on earnings and SES increases approximately as a logarithmic function of the size of the establishment for which they work. Indeed, we find zero-order correlations between schooling effects and log establishment size to be between +.88 and +.95, and statistically significant at the .025 level or better. Implications of our research are discussed.

Employers play a crucial role in the process of social stratification in the contemporary United States. The jobs they provide are the primary mechanism by which individuals are distributed among occupations and by which earnings are distributed among persons.¹ But, in spite of this fact, employers have been neglected conspicuously in the vast majority of empirical sociological research on oc-

cupational attainment and earnings in the last two decades. For example, employers (or even their aggregations, industries) do not find their way into such benchmark studies of socioeconomic achievement as the Blau-Duncan model and its extensions (Blau and Duncan, 1967; Duncan et al., 1972), the Wisconsin Model (Sewell and Hauser, 1972; 1975), or the analysis of Jencks et al. (1972) of the determinants of occupational attainment and earnings. And the few social stratification studies which do consider employers explicitly (e.g., Talbert and Bose, 1977; Bridges and Berk, 1974; 1978) analyze data on such a narrow range of occupations, such limited types of employers and such small geographic areas as to be only suggestive of general patterns by which employer characteristics fit into the process of socioeconomic achievement in the United States.²

* Address all communications to: Ross M. Stolzenberg; Department of Sociology; University of Illinois; Urbana, IL 61801.

Thanks to Ellen S. Brondfield for unflagging research assistance, to Howard Aldrich, William Form, Robert M. Hauser, Michael Moch, Reeve Vanneman, and Linda Waite for useful advice, and to the National Science Foundation for essential support (Grant SOC-77-21265).

¹ In 1974, 93.1% of the nonagricultural labor force which worked for pay in the U.S. was not self-employed (computed from U.S. Bureau of the Census, 1975:360).

² There is, however, a long tradition in sociology

Similarly, economists have done little empirical research which investigates the roles of both employer and employee characteristics in the determination of earnings. For example, Mincer's (1970) well-known review of earnings celebrates the empirical success of an earnings model—the human capital model—which completely neglects employers, and the enormous economic literature on industrial wage effects overwhelmingly is composed of research which excludes consideration of such basic worker characteristics as schooling and occupation (see Stolzenberg's [1975a] review of this literature).³

Undoubtedly, there were once good reasons for neglecting employers in the analysis of socioeconomic achievement—necessary data were often unobtainable and there were even more fundamental concerns in need of attention. But those reasons no longer hold, and the time seems right to merge thinking about employers with empirical and theoretical research on individuals' occupational attainment and earnings. Therefore, the purpose of this paper is to begin this merger in a fundamental way, at both theoretical and empirical levels, and with

of empirical research which points to the importance of employer characteristics in the process of socioeconomic achievement, but which does not specify or explain in any detail the nature of these employer effects. The most elegant analyses in this tradition are by Hauser and Featherman (1977:chap. 9), but other examples include work by Lane (1968), Blau and Duncan (1967:37) and research on the so-called situs dimension in stratification, which traces at least as far back as Morris and Murphy (1959) and Benoit-Smullyan (1944). See also the unpublished theses of Hogan (1973), Lane (1972), and Horan (1972).

³ There are some exceptions to this generalization. For example, Rees and Shultz (1970) included measures of employer characteristics in their analysis of wages in Chicago. But their study was limited to 12 blue-collar occupations, was biased by enormous nonresponse rates (56% refusals among establishments with 50 to 249 employees, for example), and is marred by the dubious practice of not presenting statistical results which differed too markedly from the results they expected before the study was done. So while the Rees and Shultz study includes both worker and employer characteristics as determinants of wages, it is more to be noted as a pioneering effort than as research which answers many questions about the process of wage determination.

data and hypotheses which pertain to the widest possible range of phenomena. In the following pages, we review and combine sociological work on organizational structure, sociological studies of social stratification, and economic research on labor market characteristics, industrial wage differentials and human capital. In the course of this review, we hypothesize that the size of an employer organization works indirectly through other dimensions of organizational structure to alter the effect of workers' schooling on their earnings and occupational attainment. Since the relationship between schooling and occupational attainment is a fundamental feature of social stratification research, and because the effect of schooling on earnings is a fundamental part of both sociological and human capital economic research on earnings, our hypothesis links organizational structure to the very heart of current issues in social stratification and human capital studies. To measure the strength of this link and to test our reasoning, we analyze data from a national survey of employed persons. Available data do not allow us to replicate earlier research on the relationship between organizational size and the dimensions of organizational structure, but that earlier work does allow us to hypothesize that schooling has a stronger effect on occupational achievement and on earnings in larger organizations than in smaller ones. On the basis of that work, we also are able to hypothesize the specific mathematical form of the relationship between organizational size and the effects of employees' schooling on their earnings and occupational status. Our empirical findings conform closely to the expected mathematical form. Finally, we discuss the implications of our ideas and findings for future work in social stratification.

I: ORGANIZATIONAL SIZE AND RETURNS TO SCHOOLING: INFERENCES FROM ORGANIZATIONAL RESEARCH

Starting with Weber's (1958:211) analysis of bureaucracy and continuing into the present, research and theory repeatedly have suggested that the structure

of organizations is strongly and causally related to the number of people which they comprise. For example, Blau and Schoenherr (1971) find organizational size to play a central role—perhaps *the* central role—in determining the structure of state employment agencies. Blau (1973) finds the number of employees in a university or college to have pervasive, large effects on its organization of academic work. A well-known study by Pugh et al. (1969:112) concludes that organizational “size causes [organizational] structuring.” Mason Haire (1959) finds very close correspondence between the growth of organizations and certain transformations in their structure. And a particularly clear and rigorous picture of the link between organizational size and structure is given by Child’s (1973) review of theoretical arguments and reanalysis of data from earlier studies. Child identifies six dimensions of organizational structure which were earlier identified in theoretical and empirical work by Pugh et al. (1968) and which correspond to measures of structure used in Blau and Schoenherr’s (1971) research.⁴ Drawing on data from five studies of organizations in the U.S., Child then shows that the correlation between the logarithm of organizational size and measures of six dimensions of organizational structure are both substantial (the mean correlation of .68) and consistent across all five data sets used.⁵ Child

also shows empirically that although “others such as Woodward (1965), J. Thompson (1967) and Aldrich (1972) have argued on theoretical and empirical grounds that the tasks and technology of an organization are more salient influences on structure than its size” (p.168), the impact of size on organizational structure remains very strong (and much stronger than task and technology variables) when the effects of task and technology variables are partialled out (p.179 and Table 8). And, finally, Child replicates and extends Blau and Schoenherr’s (1971) finding that the relationship between organizational size and the dimensions of organizational structure is approximately logarithmic. So there seems to be a solid basis for believing that the size of an organization has substantial effects on the various dimensions of its structure, and that this relationship is approximately logarithmic.

For present purposes, two dimensions of organizational structure which were identified by Pugh et al. (1968), Blau and Schoenherr (1971) and Child (1973) are particularly important. These dimensions are the extent to which procedures are standardized (hereafter called *standardization*), and the extent to which rules, procedures, instructions and communications are written (hereafter called *documentation*). Child is interested in the strong effects of organizational size on standardization and documentation, and Pugh and his colleagues are interested in the consequences of standardization and documentation for organizations. But it seems to us that these dimensions also have important effects on the socioeconomic achievement of the employees of organizations, and that these effects could very well produce a relationship between the size of employer organizations and the socioeconomic achievement of their employees. In particular, we reason that the greater the extent of documentation in an organization, (a) the greater the extent to which the work performance of its employees depends upon their ability to communicate in writing, and consequently, (b) the greater the extent to which the ability to communicate in writing affects employees’ remuneration and

⁴ These six dimensions are essentially Weberian and are supported empirically as well as being indicated theoretically (see Pugh et al., 1968, for details). While the names of the dimensions identified by Child and Pugh et al. are different from the organizational variables used by Blau and Schoenherr (1971) and by Blau (1973), five of the six dimensions correspond closely to structural properties of organizations measured (and used) in both the Blau studies. Child (1973:Table 2) shows that Pugh et al.’s dimensions have about the same correlation with organizational size as the corresponding measures of Blau. The six dimensions are: standardization, formalization, centralization, configuration, overall role specialization, and functional specialization.

⁵ The mean absolute size of the zero-order correlations between the log of organizational size and each of the six dimensions in the four studies is .68. All correlations were positive in all studies, except in the case of the dimension *overall centralization* which had negative correlations with the log of organizational size in all studies.

selection for desirable kinds of work, other things being equal. Since workers' years of schooling is widely (and at least somewhat validly) used as an indicator of their ability to communicate in writing,⁶ we also expect that the greater the extent of documentation in an organization, the greater the effect of workers' length of schooling on their earnings and occupational achievement, *ceteris paribus*. And, finally, since the extent of documentation in an organization is largely determined by the logarithm of the organization's size, we hypothesize that the effect of workers' length of schooling on their earnings and occupational attainment varies according to the logarithm of the size of the organization which employs them, *ceteris paribus*.

Our reasoning about organizational standardization leads to similar conclusions: The greater the extent to which hiring and remuneration decisions in an organization are standardized, the greater the extent to which these decisions must be based on data which lend themselves to standardized procedures. Since years of schooling is widely accepted in the U.S. as a standard (if crude) measure of cognitive ability, and since information about employees' years of schooling is available to the employer at no cost, employees' years of schooling is particularly well-suited for use in procedures concerning hiring and remuneration. Therefore, we hypothesize that the greater the extent of standardization in an organization, the greater the effect of years of its employees' schooling on their earnings and occupational achievement, other things being equal. And since standardization is strongly affected by the logarithm of organizational size, we hypothesize that the effect of workers' length of schooling on their earnings and occupational attainment varies according to the logarithm of the size of the organization which employs them, *ceteris paribus*.

A pejorative version of our hypothesis about the effects of standardization arises

from the notions of *bureaucratic* personalities discussed by Merton (1957) and *bureautic* and *bureaupathic* behavior mentioned by V. Thompson (1961:154). These supposedly all-too-common aberrations are characterized by excessive concern for rules and procedures, and are held to be the unintended results of efforts to use standardized procedures to assure that employees act efficiently and in the interest of their employers. In personnel matters, *bureaupathic* behavior would be characterized by an excessive concern for easily-measured characteristics of workers, such as their years of schooling, and a neglect of more relevant, harder-to-observe criteria. Indeed, some authors have argued that schooling is only remotely related to performance on most jobs, and that the widespread use of schooling in personnel decisions is mere *credentialism* (Miller, 1968) which constitutes a "great training robbery" (Berg, 1970).⁷ Thus, since *bureaucratic* personalities and *bureaupathic* behavior are caused by standardization, and since these aberrations are held to increase the effects of workers' schooling on their occupational achievement and earnings, this literature leads to the same hypothesis that we drew from a less pejorative view of standardization: Other things being equal, the greater the extent of standardization in an organization, the greater the effect of its employees' years of schooling on their earnings and occupational achievement. And since standardization is strongly determined by the logarithm of organizational size, we once again hypothesize that the effect of workers' length of schooling on their earnings and occupational attainment varies according to the logarithm of the size of the organization which employs them, *ceteris paribus*.

So we have offered three separate mechanisms by which an organization's size may alter the effects of its employees' schooling on their socioeconomic achievement. We now consider some counterarguments.

⁶ In recent years there have been extensive arguments that workers' schooling is falsely believed to be related to their work performance. We consider such arguments later in this paper.

⁷ Related issues have been treated at the societal and institutional levels of analysis. See, in particular, Meyer (1977) and Collins (1977).

II: DOESN'T THE LABOR MARKET HOMOGENIZE THE PROCESS OF SOCIOECONOMIC ACHIEVEMENT?

The most serious challenge to our reasoning in the previous section comes from neoclassical economics. Neoclassical theory argues that the *invisible hand* of the market homogenizes the wage determination process, preventing workers from earning more with one employer than they would be paid by another employer. Neoclassical theory is not without problems,⁸ but we think that it raises an important question which we must confront: How can the processes of wage determination and occupational attainment be different in some organizations than in others? Without launching a wholesale attack on the neoclassical perspective, we offer the following answer to that question, based on recent work on the segmented economy and on internal labor markets: Economic markets, including labor markets, are partitioned into segments. Circumstances, institutions and market mechanisms vary from one segment to another. In particular, large and small organizations tend to be located in different market segments, thereby allowing market processes affecting socioeconomic attainment to operate differently in small organizations than in large ones. We now review the *segmented economy* and *internal labor market* perspectives in more detail.

The keystone of segmented economy theory is Averitt's (1968) argument that there are two types of firms: core firms and periphery firms.⁹ Each core firm is diversified geographically, sells a variety of products, has extensive assets, borrows money at preferential interest rates, and can expand to become its own supplier or distributor, if necessary. As a result, core firms are buffered against economic fluc-

tuations, can grow into profitable new areas and can survive temporary setbacks. Periphery firms are the opposite: because they are not diversified, economic fluctuations in one product or geographic area can easily prove disastrous for them. And because they have small assets, periphery firms have difficulty borrowing the resources to survive setbacks, to improve their operations or to expand into profitable areas. In brief, periphery firms are small and weak, and their condition tends to make them smaller and weaker over time; core firms are big and strong, and they tend to become bigger and stronger over time. Periphery firms try to cope with their weak economic condition by paying low wages, offering small nonwage benefits, providing few on-the-job amenities, and laying off or firing workers during slack periods. Further, the small size, slow or nonexistent growth, and economic frailty of periphery firms combine to offer their employees only limited possibilities to advance to better jobs within the firm. In contrast, core firms are able to offer high wages, benefits and amenities, as well as steadier work and more within-firm advancement opportunities. Thus, segmented economy theory offers a specific linkage between firm size and the mechanisms governing employee socioeconomic achievement.

Closely related to the segmented economy perspective just discussed are internal labor market theories (e.g., Doeringer and Piore, 1971; Althauser and Kalleberg, 1977; and especially the empirical study by Alexander, 1974).¹⁰ Internal market theories combine the notion of segmented economy with Kerr's (1954) idea of labor market *balkanization*, or partition. Akin to Kerr's *manorial* market type, the internal market exists entirely within a single firm. Workers enter the internal market through "port-of-entry" jobs and can reach higher positions only by promotion from entry-level jobs. Two points about internal labor markets are relevant to

⁸ See Gordon (1972) or Butler (1961) for a discussion of these difficulties.

⁹ The dual economy model has been expanded into a tripartite schema (Bluestone, 1970), applied to the manpower problems of blacks (Victorisz and Harrison, 1970; Leigh, 1976), elaborated into a general statement of labor market fractures (Reich et al., 1973) and reviewed from both neoclassical and radical perspectives (Cain, 1975; 1976; Gordon, 1972).

¹⁰ For a discussion of the relationship between dual labor market theory and internal labor market theories, see Gordon (1972:15-6).

present purposes. First, internal markets can exist only in large organizations—it takes more than a handful of jobs and job-holders to establish a market. And, second, except at entry-level positions, the processes of wage-setting and occupational achievement in the internal market are shielded somewhat from the forces of supply, demand and social custom which govern wage-setting and occupational achievement in the external market. Thus, internal labor market theories, like the segmented economy perspective, offer an explicit mechanism by which the processes of socioeconomic achievement can be different in large organizations than in small or medium-sized organizations.

III: EMPIRICAL ANALYSIS

Our discussion in the previous pages leads us to investigate two hypotheses: First, the hypothesis of earlier researchers that, *ceteris paribus*, the earnings of employees increase with the size of the organization which employs them; and, second, our own hypothesis that the effect of workers' length of schooling on their earnings and occupational attainment varies with the logarithm of the size of the organization which employs them. To test these hypotheses, we first divide a sample of employed persons into subgroups according to the size of the organization in which they are employed. Then, separately in each size-specific subgroup, we estimate a model of earnings and occupational attainment. From the parameters of these employer size-specific models, we calculate the effect of schooling on earnings and on occupational attainment for employees of different-sized employers. And by graphic and statistical analysis, we determine if these effects are a logarithmic function of employer size, as hypothesized. Similarly, by examining the parameters of these employer size-specific models, we determine if employees of smaller organizations have lower earnings than employees of larger organizations, net of other factors. We now describe the data we use in these analyses, then describe our models of earnings and occupational attainment, and our findings.

A. *The Data*

Data are drawn from the 1973 Quality of Employment Survey conducted by the Survey Research Center of the University of Michigan, which describes the sample as follows:

Data for this study were obtained through personal interviews with 1,496 persons living in housing units within the United States and the District of Columbia, exclusive of households on military reservations, and households in Hawaii and Alaska. Eligible persons were household members 16 years of age or older who were employed for pay for 20 or more hours per week. People were also interviewed if they worked for pay but were currently not working due to strike, sickness, weather, vacation, or for personal reasons

Approximately 70 percent of the households had one or more persons who met the eligibility criteria for respondent selection, of which 1,982 persons were designated as respondents; of these 75.5 percent were interviewed. (Survey Research Center, 1975:IV)

The sample design for this study was the usual multistage probability sample of the Survey Research Center. Data were weighted to reflect sampling probabilities.

Farmers, farm managers and farm laborers were excluded from our analyses to avoid ambiguities in earnings data caused by payment in kind. Blacks and women also were excluded because they were not represented in sufficient numbers to allow separate analyses by sex and race. And self-employed persons were excluded to avoid ambiguities in distinguishing between earnings and income from business investments. Thus, our sample was reduced to 690 respondents.

Employer size was measured by the respondent's answer to the query, "About how many people work at the location where you work—I mean all types of workers in all areas and departments?" We divided responses to this question into five different categories: 1 to 9 persons, 10 to 49, 50 to 99, 100 to 499, and 500 or more persons. These five categories defined the five size groups used in our analysis. (See Table 1 for size-specific *n*'s.)

Our employer size measure (which is the only employer size measure in the

Table 1. Effects of Schooling on Occupational Status and Annual Earnings, by Establishment Size, with Related Statistics

Establishment Size (number of employees)	Data Cases ¹ A	Effects of Schooling on Occupational Status			Effects of Schooling on Earnings	
		Based on Equation (1)		Based on Equation (2)	Based on Equation (3)	Based on Equation (4)
		Metric Effects ² B	Standardized Effects C	Metric Effects ² D	Rates of Return (%) ³ E	Rates of Return (%) ³ F
1-9	117	3.68	.508	3.72	0.90%	-1.19%
10-49	178	4.33	.585	4.58	3.13	-0.02
50-99	73	4.21	.658	4.23	4.96	1.46
100-499	152	4.55	.692	4.89	4.48	1.88
≥500	170	5.27	.695	5.64	5.00	1.41
Correlation with Log _e (Midpoint of Size Category) ⁴	...	+.9446	+.9542	+.9299	+.8821	+.8811

¹ Actual number of data cases used in regressions and significance tests is slightly smaller, due to missing data.

² Duncan SEI points per year of schooling.

³ Percent increase in annual earnings per additional year of schooling.

⁴ Pearsonian correlation of log_e (midpoint of size category) with effect of schooling. See text for details of computation and significance tests.

Quality of Employment Survey) is an indicator of *establishment* size. We would have liked to have used firm size in addition to establishment size, for some large organizations are composed of many small establishments. However, we have been unable to discover *any* individual-level data based on a sample of the U.S. population which includes both establishment and employer size, as well as individuals' schooling, earnings and other variables which are essential for our analyses. In any case, we note that small employer size is a sufficient condition for small establishment size, and large establishment size is a sufficient condition for large employer size, highlighting the clear connection between employer and establishment size.

B. Employer Size and Occupational Status

We start with a stark model of occupational attainment. In this first model, occupational status (Occ) is caused by workers' length of schooling (Ed) and length of labor market experience (Ex).

Ed and Ex are measured in years; Occ is measured on Duncan's socioeconomic index for occupations. Labor force experience is held constant in this model because occupational status has been shown to increase with workers' time in the labor force (Duncan et al., 1972:263). Mathematically, this model is given in equation (1) and its parameters are estimated by regression analysis:

$$\widehat{\text{Occ}} = a + b_1 \text{Ed} + b_2 \text{Ex}. \quad (1)$$

The metric (unstandardized) coefficient for Ed indicates the average number of occupational status points produced by an additional year of schooling.

We estimated equation (1) separately in each of the five size groups identified earlier. The coefficient for Ed in each of these size-specific analyses is shown in the far left column of Table 1, and we have plotted these coefficients against the logarithm of establishment size in Figure 1. Scanning down column A of Table 1, notice the large impact of employer size on the effect of schooling: the coefficient

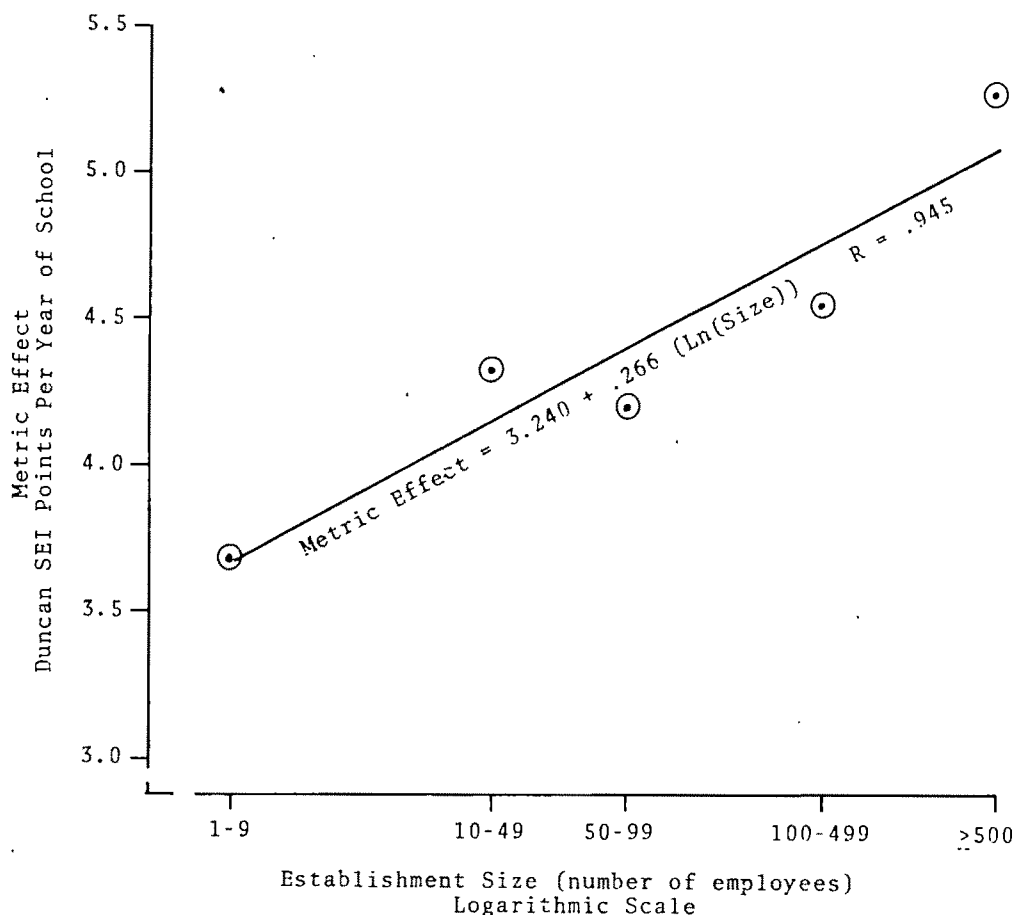


Figure 1. Metric Effect of Schooling on Occupational SES, by Establishment Size

for Ed is 43% larger in the biggest size group than in the smallest group ($5.266/3.677=1.43$). Looking at Figure 1, notice that the effect of schooling on occupational attainment conforms to the relationship we hypothesized. More rigorously, we correlated the logs of the midpoints of the size categories with the effect of schooling on occupational attainment for workers employed by organizations in those categories.¹¹ That correlation was $+.94$, indicating a rather close fit between the log of establishment size and the effect of schooling on occupational status.¹²

¹¹ We estimate an average establishment size of 1,000 for the 500-and-over size category.

¹² Significant at the .025 level, one-tailed test (direction predicted from the substantive hypothesis under consideration). An analysis of covariance rejects the null hypothesis that the coefficients in equation (1) are the same in all five size categories (at a

Since it is common practice to use standardized (beta) coefficients to measure the effect of schooling on occupational status, column 2 of Table 1 displays the standardized form of the coefficient for Ed in equation (1), and Figure 2 plots these standardized effects against the logarithm of establishment size. Similar to results obtained for metric coefficients, the beta of Ed is 37% larger in the biggest size group than in the smallest category, and the correlation between the standardized effects and the log of the midpoint of the size categories is $+.95$.¹³ Thus, our stark model of occupational attainment produces results which are thoroughly consistent with our hypothesis about the rela-

significance level of .01). See Johnston (1972:192-207) for details of covariance analysis calculations.

¹³ Significant at the .01 level (one-tailed test).

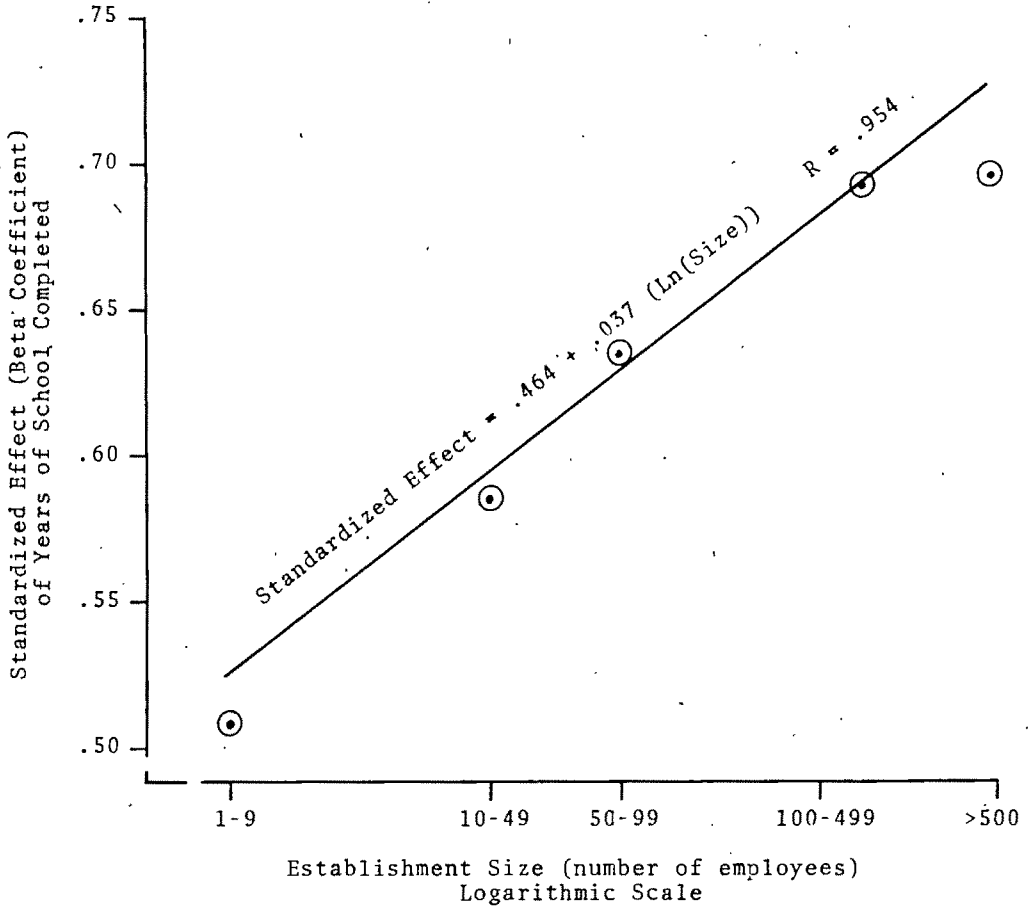


Figure 2. Standardized Effects of Schooling on Occupational SES, by Establishment Size

tionship between employer size and the effect of schooling on occupational achievement. We now turn to a less stark (and therefore more realistic) model.

Equation (2) is similar to our first model, except that it includes the respondent's years of tenure with his current employer (Ten), and it also includes a multiplicative interaction between Ed and Ex. Including tenure in the model controls for the possibility that employees of larger organizations are less prone to change employers than employees of smaller organizations, and that their longer tenure affects their probability of being promoted into higher status occupations. Including the interaction between Ed and Ex in the model allows the effect of schooling to be stronger at some stages of the career than at others, and is therefore consistent with findings of Blau and Duncan (1968:170;

this inference is drawn from computations of indirect effects in the Blau-Duncan *Basic Model*). By including the Ed•Ex interaction, we avoid being misled by a possible relationship between employer size and workers' length of experience.

$$\widehat{Occ} = a + b_1 Ed + b_2 Ex + b_3 Ed \cdot Ex + b_4 Ten. \quad (2)$$

Because schooling appears in two terms in equation (2), its effect on occupational status must be measured by calculating the partial derivative of Occ with respect to Ed, $\partial Occ / \partial Ed$. The partial derivative is directly comparable to the metric regression coefficient: $\partial Occ / \partial Ed$ is the rate of change in occupational prestige per unit of change in schooling.¹⁴

¹⁴ The formula for the partial derivative $\partial Occ / \partial Ed$

Results of calculations based on equation (2) are presented in column C of Table 1. Once again, these findings are entirely consistent with our hypothesis: The effect of Ed on Occ is 52% larger in the biggest size group than in the smallest size category. And the correlation between the effect of Ed and the logs of the midpoints of the size categories is +.93.¹⁵

So, to sum up briefly, we have found that the effect of workers' schooling on their occupational status varies very strongly with the logarithm of the size of the establishment which employs them. We also have found that these employer-size effects are substantial, and that the effect of schooling on occupational status is about 40 to 50% higher in establishments with more than 500 employees than in establishments with less than 10 employees. We now turn to the effects of employer size on the relationship between workers' schooling and earnings.

C. Employer Size and Effects of Schooling on Earnings

Equation (3) presents a model of earnings. The dependent variable in this model (\$) is the natural logarithm of the respondent's annual earnings. *Urban*, *West*, and *South* are dummy variables which indicate whether or not the respondent is a resident of an urban area, the West, or the South, respectively.¹⁶ Ed, Ex and Ten are the same variables used in our analysis of occupational SES; they measure schooling, length of labor force experience and length of employment with the current employer. We use the logarithm of earnings in equation (3) because there is overwhelming evidence that the earnings function is logarithmic in form (see, e.g., Heckman and Polachek, 1974; Griliches, 1977), and because the distribution of

earnings is known to be more like the log normal distribution than the normal distribution (see Hill, 1959; Cramer, 1971:68; Stolzenberg, 1975b:652). Dummy variables for geographic location are included to account for geographic variation in money wages (as opposed to real wages; see Block, 1948; Kerr, 1957; Galloway, 1963; Folger and Nam, 1967; Reeder, 1957). Ed² and Ed•Ex are included to account for nonlinearities and nonadditivities in the effects of schooling on earnings (see Heckman and Polachek, 1974; Thurow, 1967; Stolzenberg, 1975a). Occupation is not included in this model, but occupational status is included in later models when we consider its role as a variable which intervenes between schooling and earnings in the wage determination process.

$$\hat{\$} = a + b_1 \text{ Urban} + b_2 \text{ West} + b_3 \text{ South} + b_4 \text{ Ed} + b_5 \text{ Ed}^2 + b_6 \text{ Ex} + b_7 \text{ Ed} \cdot \text{Ex} + b_8 \text{ Ten.} \quad (3)$$

As in the case of our second model of occupational status, nonlinearities and nonadditivities in equation (3) require the use of partial derivatives to measure the effect of schooling on earnings. But because the dependent variable in equation (3) is the natural logarithm of annual earnings, the partial derivative $\partial\$/\partial\text{Ed}$ give the *proportional* change in earnings per addition year of schooling.¹⁷ Thus, the partial derivative is a rate of return and can be compared with other rates of return (such

¹⁷ To see this, let Y = annual earnings. Thus, $Y = e^{\$}$, and $Y = e^{a + b_1 \text{ Urban} + b_2 \text{ West} + b_3 \text{ South} + b_4 \text{ Ed} + b_5 \text{ Ed}^2 + b_6 \text{ Ex} + b_7 \text{ Ed} \cdot \text{Ex} + b_8 \text{ Ten} + \epsilon}$, where ϵ is an error term. Differentiating,

$$\partial Y / \partial \text{Ed} = (e^{a + b_1 \text{ Urban} + b_2 \text{ West} + b_3 \text{ South} + b_4 \text{ Ed} + b_5 \text{ Ed}^2 + b_6 \text{ Ex} + b_7 \text{ Ed} \cdot \text{Ex} + b_8 \text{ Ten} + \epsilon}) (b_4 + 2b_5 \text{ Ed} + b_7 \text{ Ex}).$$

is $b_1 + b_5 \text{ Ex}$. Thus, the effect of Ed on Occ varies according to the value of length of experience. We evaluated the partial derivative at 19 years of experience, which is approximately the mean of Ex in the data we are using.

¹⁵ Significant at the .025 level (one-tailed test).

¹⁶ West and South as used here conform to definitions by the U.S. Census Bureau. We use the term *Urban* to mean a Standard Metropolitan Statistical Area and its suburban ring.

Now, $\partial\$/\partial\text{Ed} = b_4 + 2b_5 \text{ Ed} + b_7 \text{ Ex}$, so, substituting terms, $\partial Y / \partial \text{Ed} = Y(\partial\$/\partial\text{Ed})$. Dividing

through by Y, $\frac{\partial Y / Y}{\partial \text{Ed}} = \partial\$/\partial\text{Ed}$. And since $\partial Y / Y$ is the

proportional change in annual earnings, $\partial\$/\partial\text{Ed}$ gives the proportional rate of change in earnings per unit change in Ed. We evaluated the partial derivatives at the approximate sample means of Ex (19 years) and Ed (12 years).

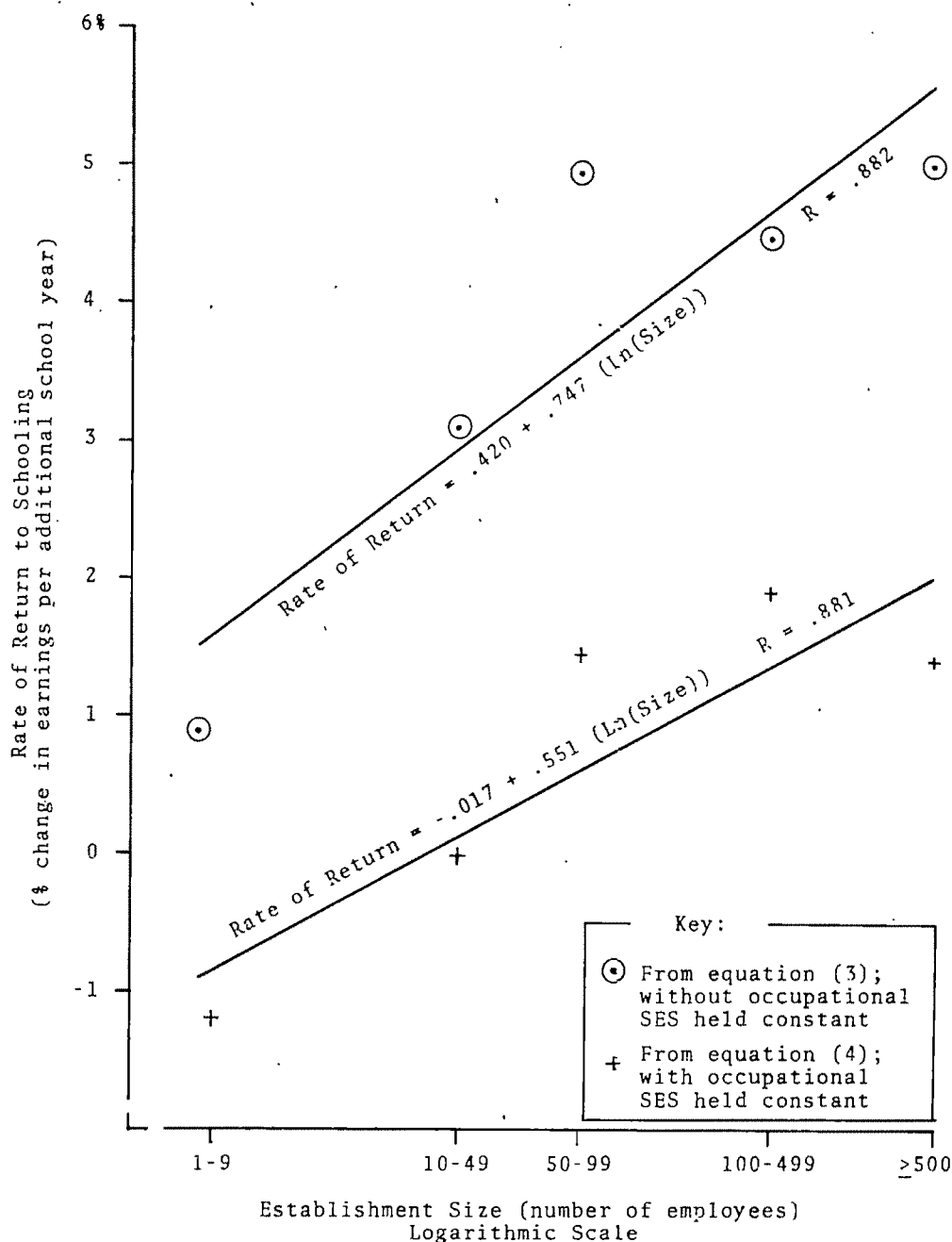


Figure 3. Rate of Return to Schooling, by Establishment Size

as interest rates on savings accounts) to get some intuitive feel for their size.

As in our analysis of school effects on occupational status, we calculated the effect of schooling on earnings in each of the five employer size categories we defined earlier. Column E of Table 1 displays the

employer size-specific effects of schooling on earnings. Figure 3 plots those effects against the logarithms of the midpoint of the size categories. Looking at the figure or the table, notice that these findings are consistent with our hypothesis. The effect of workers' schooling on their earnings is

2.6 times as large in the biggest size group as it is in the smallest size category ($2.6 = 5.39/2.06$). Further, we found a correlation of $+ .88$ between the logs of the midpoints of the size categories and the effects of schooling on earnings in those categories.¹⁸

D. Occupation Intervenes

Consistent with our reasoning, we have found that the effect of workers' schooling on their earnings and occupational status varies with the log of the size of the establishment in which they work. But since schooling is causally prior to occupation, and since occupation is causally prior to earnings, it is appropriate to measure the extent to which the size-schooling-earnings interaction is mediated by the size-schooling-occupational status relationship. In less technical terms, we want to know if employer size has separate effects on both the schooling-earnings and the schooling-occupational status relationships. In order to answer this question, we add occupational status (Occ) as a cause of annual earnings, yielding equation (4). We proceed with equation (4) as we did with equation (3), and obtain the rates of return to schooling which are displayed in Column F of Table 1.

$$\begin{aligned} \hat{\$} = & a + b_1 \text{ Urban} + b_2 \text{ West} + b_3 \text{ South} \\ & + b_4 \text{ Ed} + b_5 \text{ Ed}^2 + b_6 \text{ Ex} \\ & + b_7 \text{ Ed} \cdot \text{Ex} + b_8 \text{ Ten} \\ & + b_9 \text{ Occ.} \end{aligned} \quad (4)$$

We also have used regression to fit the relationship between log of establishment size and the rate of return to schooling. These regression fits are shown in Figure 3. Looking at the figure, notice that while rates of return to schooling based on equation (4) are lower than rates of return based on equation (3), the slope of the line relating log establishment size to the rate of return is diminished only by about a quarter when occupational status is included in the analysis.¹⁹ And we find a

correlation of $+ .88$ between the logs of the midpoints of the employer size intervals and the rates of return to schooling within those intervals.²⁰ Thus, there appear to be substantial size-of-establishment effects on the relationship between schooling and occupational status, and there *also* appear to be separate direct effects of establishment size on monetary returns to schooling.

E. Total Effects of Employer Size on Earnings

Our findings of establishment size effects on the schooling-earnings relationship begs an important question which has occupied earlier researchers: Would a worker tend to earn more working in a large establishment than if employed in a small one? Others have hypothesized and demonstrated that employee earnings are positively associated with establishment size (see, e.g., Lester's [1967] exposition of 25 years of data on the subject; see also Reynolds [1966:271]).²¹ But previous studies of establishment size effects on earnings do not control for establishment size differentials in employee characteristics, and/or they do not allow for establishment size differences in the effects of worker characteristics on earnings (e.g., Spaeth, 1976). We now apply procedures which avoid these shortcomings of past work.

First, looking at the top row of Table 2, notice that the mean log earnings increase dramatically and monotonically with establishment size. Applying laws of logarithms, these data indicate that earnings in the largest size category are about

¹⁸ Significantly different from zero at the .025 level (one-tailed test).

¹⁹ The slope of the line based on equation (4) is .551; the slope of the line based on equation (3) is .747. $1 - .551/.747 = .26$, or about one quarter.

²⁰ Significantly different from zero at the .025 level (one-tailed test).

²¹ Some analysts have linked establishment size to worker earnings on the basis of a relationship between firm size and profitability (Lester, 1967:65; Reynolds, 1966:271). There is evidence that large firms tend to have less competition than small firms, and therefore that large firms tend to be more profitable than smaller firms (Scherer, 1970; Shepard, 1970). And there is industry-level evidence that more profitable firms pay higher wages than less profitable firms (Garbarino, 1950; Levinson, 1960; Rees, 1962; Weiss, 1966; Wachter, 1970). The Quality of Employment Survey includes no data on employer profitability.

Table 2. Observed and Standardized Mean Log. Earnings, by Establishment Size

	Establishment Size (number of employees)				
	1-9	10-49	50-99	100-499	≥500
Observed	8.916	9.096	9.155	9.243	9.306
Standardized	9.074	9.078	9.162	9.236	9.254

Note: See text for details of standardization procedure.

48% larger than earnings in the smallest size category ($1.48 = \exp [9.306 - 8.916]$). This pattern of large establishment size differentials establishes the consistency of the Quality of Employment Survey data with data used in previous studies of establishment size differentials in earnings.

To hold constant the effects of size-of-establishment variation in *worker* characteristics, we applied a regression standardization procedure popularized by Duncan (1969). We first estimated the parameters of equation (4) separately in each of the five size categories identified earlier in this paper.²² Then, separately in each size category, we used the all-sample means of the independent variables in combination with the size-specific parameters of equation (4) to estimate the mean log earnings. This procedure allowed the process of earnings determination to vary by establishment size, but it held constant the worker characteristics on which the process operated. Thus, differences between the standardized mean log earnings are indicative of establishment size effects on earnings, net of variation in worker characteristics included in equation (4). Looking at the results of this standardization in column 2 of Table 2, notice that the standardized means, like the unstandardized means, increase monotonically with establishment size. Applying the laws of logarithms to these figures indi-

cates that the standardized mean in the largest size category is about 20% larger than the standardized mean in the smallest size group ($1.20 = \exp [9.2543 - 9.0741]$). Recalling that the unstandardized data showed a differential of 48% between the largest and the smallest size groups, we can conclude that differences in levels of employee characteristics account for about three-fifths of the total differential, but the employer size effects on earnings remain quite substantial even when employee characteristics are held constant.

IV: CONCLUSIONS

Our purpose here has been to consider the role of a fundamental characteristic of employers—their size—on the processes affecting the socioeconomic achievement of their employees. In the theoretical section of this paper, we reviewed a wide body of theory and research which suggested several mechanisms by which employers' size could alter the effect of employees' schooling on their earnings and occupational status. Past research and theory led us to hypothesize that one or more of the following three mechanisms could produce higher pay and occupational status returns to schooling in large organizations than in small ones. First, school-learned skills could be related to performance of activities which tend to take place more frequently in large organizations than in small ones. Second, school-learned skills could be merely widely believed to be related to activities which take place more frequently in large organizations than in small ones; these widespread beliefs would lead large employers to reward schooling more highly than small employers. And, third, the structure of large organizations creates conditions which make it especially convenient for them to act *as if* schooling were relevant to job perform-

²² We used the analysis of covariance to test the hypothesis that the constant terms in equation (4) are the same in all five size categories and to test the hypothesis that the coefficients for independent variables are the same in all five categories. Both hypotheses were rejected at a significance level of 1%. However, as is often the case in such analyses, multicollinearity produced highly unstable estimates of the main effects of establishment size, and we therefore do not interpret these estimates. See Johnston (1972:192-207) for details of covariance analysis tests used here.

ance, whether or not a relationship between workers' schooling and productivity actually exists. Our theoretical analysis also considered the structural features of labor markets which would allow (or even foster) employer effects on the process of socioeconomic achievement.

In our empirical analysis, we found almost exactly what we hypothesized about the relationship between establishment size and the effects of schooling on occupational status and earnings: The effect of schooling on occupational status and the effect of schooling on earnings varies as a linear function of the logarithm of the size of the establishment in which a person is employed. Indeed, the zero-order correlations between log establishment size and the effects of schooling were all between $+.88$ and $+.95$, and all were statistically significant at the $.025$ level or better.

Our empirical analysis also reconsidered a wide body of literature which has ignored the effects of schooling on wages, but which has claimed nonetheless that earnings are higher in large establishments than in small ones. Although we found past studies of this subject wanting on methodological grounds, we found that correction of methodological problems reduced but by no means eliminated size-of-firm differentials in earnings.

In closing, we briefly mention some implications and consequences of the arguments and analyses presented here. First, and most important, we think that our findings, together with the findings of other more limited analyses of narrower topics and more restricted data bases, constitute a compelling argument for systematically drawing organizational research into the study of social stratification. Understanding the effect of workers' schooling on their earnings and occupational status has been a central problem in American stratification studies for at least the last 15 years, and we have shown that these schooling effects are altered substantially by a very basic characteristic of the organization in which workers are employed. Second, although we are not economists, and though our purpose here has not been to contribute to economics, we also note that our findings would seem to have important implications for certain

work in that discipline. The central empirical problem of human capital economics in the last 14 years has been measuring workers' rates of return to schooling. Virtually without exception, these measurements have been made by assuming that employer characteristics (or the *demand side*) can be ignored in short-run analyses of schooling effects on earnings. Our findings seem to contradict this assumption and suggest that even in the short run, employer characteristics have very strong effects on their employees' rates of return to schooling. And, third, the findings and arguments presented in this paper seem to have important implications for social policy and for future social trends. In particular, our analyses would support speculation that increases in the average size of establishments will tend to increase the importance of schooling in determining workers' earnings and occupational status.²³ Of course, any extrapolation from our cross-sectional analyses to long term trends is speculation which merely suggests questions to be resolved by further empirical research. But the need for further analysis is one of the central arguments of this paper. It is time to bring employers back into the sociological study of employment.

REFERENCES

- Aldrich, Howard
1972 "Technology and organizational structure: a reexamination of the findings of the Aston group." *Administrative Science Quarterly* 17:26-43.
- Alexander, Arthur J.
1974 "Income experience and the structure of internal labor markets." *Quarterly Journal of Economics* 88:63-85.
- Althausen, Robert and Arne Kalleberg
1977 "Occupational and firm labor markets: toward a specification of structural factors in wage attainment." Paper presented at the annual meeting of the American Sociological Association, Chicago.
- Averitt, Richard T.
1968 *The Dual Economy: The Dynamics of*

²³ Without pretending to perform an analysis of trends in establishment size, we note that in 1954, 32% of employees in manufacturing worked in establishments with 20 or more employees, and in 1972, 35% of employees in manufacturing worked in establishments with at least 20 employees (U.S. Bureau of the Census, 1975).

- American Industry Structure. New York: Norton.
- Benoit-Smullyan, Emile
1944 "Status, status types, and status interactions." *American Sociological Review* 9:151-61.
- Berg, Ivar
1970 *Education and Jobs: The Great Training Robbery*. New York: Praeger.
- Blau, Peter
1973 *The Organization of Academic Work*. New York: Wiley.
- Blau, Peter and Otis Dudley Duncan
1967 *The American Occupational Structure*. New York: Wiley.
- Blau, Peter and Richard A. Schoenherr
1971 *The Structure of Organizations*. New York: Basic Books.
- Block, Joseph W.
1948 "Regional wage differentials 1907-1946." *Monthly Labor Review* 66:371-7.
- Bluestone, Barry
1970 "The tripartite economy: labor markets and the working poor." *Poverty and Human Resources* 5:15-36.
- Bridges, William P. and Richard A. Berk
1974 "Determinants of white-collar income: an evaluation of equal pay for equal work." *Social Science Research* 3:211-33.
1978 "Sex, earnings, and the nature of work: a job-level analysis of male-female income differences." *Social Science Quarterly* 58:553-65.
- Butler, Arthur D.
1961 *Labor Economics and Institutions*. New York: Macmillan.
- Cain, Glenn
1975 "The challenge of dual and radical theories of the labor market to orthodox theory." *American Economic Review* 65:16-22.
1976 "The challenge of segmented labor market theories to orthodox theory: a survey." *Journal of Economic Literature* 14:1215-57.
- Child, John
1973 "Predicting and understanding organization structure." *Administrative Science Quarterly* 19:168-85.
- Collins, Randall
1977 "Some comparative principles of educational stratification." *Harvard Educational Review* 47:1-27.
- Cramer, J. S.
1971 *Empirical Econometrics*. New York: Elsevier.
- Doeringer, Peter and Michael Piore
1971 *Internal Labor Markets and Manpower Analysis*. Lexington: Lexington Books.
- Duncan, O. D.
1969 "Inheritance of poverty or inheritance of race." Pp. 85-110 in D. P. Moynihan (ed.), *Understanding Poverty*. New York: Basic Books.
- Duncan, Otis Dudley, David L. Featherman and Beverly Duncan
1972 *Socioeconomic Background and Achievement*. New York: Seminar Press.
- Folger, J. K. and Charles P. Nam
1967 *Education of the American Population*. Washington, D.C.: U.S. Government Printing Office
- Galloway, L.
1963 "The North-South wage differential." *Review of Economics and Statistics* 45:264-72.
- Garbarino, Joseph W.
1950 "A theory of interindustry wage structure variation." *Quarterly Journal of Economics* 64:287-305.
- Gordon, David M.
1972 *Theories of Poverty and Underemployment: Orthodox, Radical, and Dual Labor Market Perspectives*. Lexington: Lexington Books.
- Griliches, Zvi
1977 "Estimating the returns to schooling: some econometric problems." *Econometrica* 45:1-22.
- Haire, Mason
1959 "Biological models and empirical histories of the growth of organizations." Pp. 287-93 in M. Haire (ed.), *Modern Organization Theory*. New York: Wiley.
- Hauser, Robert M. and David L. Featherman
1977 *The Process of Stratification*. New York: Academic Press.
- Heckman, J. and S. Polachek
1974 "Empirical evidence of the functional form of the earnings-schooling relationship." *Journal of the American Statistical Association* 69:350-4.
- Hill, T. P.
1959 "An analysis of wages and salaries in Great Britain." *Econometrica* 27:355-76.
- Hogan, Dennis P.
1973 *Industrial Mobility and the Situs Effects of Industry*. Master's thesis, Department of Sociology, University of Wisconsin, Madison.
- Horan, Patrick M.
1972 *The Structure of Occupational Mobility: A Comparative Approach*. Ph.D. dissertation, Department of Sociology, University of Wisconsin, Madison.
- Jencks, Christopher, Marshall Smith, Henry Acland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns and Stephen Michelson
1972 *Inequality*. New York: Basic Books.
- Johnston, J.
1972 *Econometric Methods*. 2nd ed. New York: McGraw-Hill.
- Kerr, Clark
1954 "The balkanization of labor markets." Pp. 92-110 in E. Wright Bakke, P. M. Hauser, G. L. Palmer, C. A. Myers, D. Yoder, C. Kerr (eds.), *Labor Mobility and Economic Opportunity*. Cambridge, Ma.: Technology Press; New York: Wiley.
1957 "Wage relationships: the comparative impact of market and power forces." Pp. 173-93 in J. T. Dunlop (eds.), *The Theory of Wage Determination*. London: Macmillan.



- Lane, Angela
1968 "Occupational mobility in six cities." *American Sociological Review* 33:740-9.
- 1972 *Contexts of Socioeconomic Attainment: The Role of Community, Industry of Employment and Spatial Mobility*. Ph.D. dissertation, Department of Sociology, University of Chicago.
- Leigh, Duane E.
1976 "Occupational achievement in the late 1960s: an indirect test of the dual labor market." *Journal of Human Resources* 2:155-71.
- Lester, Richard
1967 "Pay differentials by size of establishment." *Industrial Relations* 7:57-67.
- Levinson, Harold M.
1960 "Postwar movements of prices and wages in manufacturing industries." Study paper no. 21, Joint Economic Committee, 86th Congress, 2nd Session. Washington, D.C.: U.S. Government Printing Office.
- Merton, Robert K.
1957 *Social Theory and Social Structure*. New York: Free Press.
- Meyer, John W.
1977 "The effects of education as an institution." *American Journal of Sociology* 83:55-77.
- Miller, S. M.
1968 *Breaking the Credentials Barrier*. New York: Ford Foundation.
- Mincer, Jacob
1970 "The distribution of labor incomes: a survey." *Journal of Economic Literature* 8:1-26.
- Morris, Richard T. and R. J. Murphy
1959 "The situs dimension in occupational structure." *American Sociological Review* 24:231-9.
- Pugh, D. S., D. J. Hickson, C. R. Hinings and C. Turner
1968 "Dimensions of organization structure." *Administrative Science Quarterly* 13:65-91.
- 1969 "The context of organizational structure." *Administrative Science Quarterly* 14:91-114.
- Reeder, M.
1957 *Labor in a Growing Economy*. New York: Wiley.
- Rees, Albert
1962 *New Measures of Wage-Earner Compensation in Manufacturing*. New York: National Bureau of Economic Research.
- Rees, Albert and George P. Shultz
1970 *Workers and Wages in an Urban Labor Market*. Chicago: University of Chicago Press.
- Reich, Michael, David Gordon and Richard Edwards
1973 "Theory of labor market segmentation." *American Economic Review, Papers and Proceedings* 63:359-65.
- Reynolds, Lloyd G.
1966 *Economics: A General Introduction*. Homewood: Irwin.
- Scherer, F. M.
1970 *Industrial Market Structure and Economic Performance*. Chicago: Rand McNally.
- Sewell, William H. and Robert M. Hauser
1972 "Causes and consequences of higher education: models of the status attainment process." *American Journal of Agricultural Economics* 54:851-61.
- 1975 *Education, Occupation and Earnings: Achievement in the Early Career*. New York: Academic Press.
- Shepard, W. G.
1970 *Market Power and Economic Welfare*. New York: Random House.
- Spaeth, Joe L.
1976 "Characteristics of the work setting and the job as determinants of income." Pp. 161-76 in W. H. Sewell, R. M. Hauser and D. L. Featherman (eds.), *Schooling and Achievement in American Society*. New York: Academic Press.
- Stolzenberg, Ross M.
1975a "Education, occupation, and wage differences between white and black men." *American Journal of Sociology* 81:299-323.
- 1975b "Occupations, labor markets and the process of wage attainment." *American Sociological Review* 40:645-65.
- Survey Research Center
1975 *The 1972-1973 Quality of Employment Survey*. Institute for Social Research, University of Michigan, Ann Arbor.
- Talbert, Joan and Christine E. Bose
1977 "Wage attainment processes: the retail clerk case." *American Journal of Sociology* 83:403-24.
- Thompson, James D.
1967 *Organizations in Action*. New York: McGraw-Hill.
- Thompson, Victor
1961 *Modern Organizations*. New York: Knopf.
- Thurow, L.
1967 *The Occupational Distribution of the Returns to Education and Experience for Whites and Negroes*. Proceedings of the Social Statistics Section of the American Sociological Association.
- U.S. Bureau of the Census
1975 *Statistical Abstract of the United States: 1975*. Washington, D.C.: U.S. Government Printing Office.
- Vietorisz, Thomas and Bennett Harrison
1970 *The Economic Development of Harlem*. New York: Praeger.
- Wachter, Michael L.
1970 *Relative Wage Determination among Industries*. Ph.D. dissertation, Department of Economics, Harvard University.
- Weber, Max
1958 *From Max Weber: Essays in Sociology*. Ed. by Gerth, H. H. and C. Wright Mills. New York: Oxford University Press.
- Weiss, Leonard W.
1966 "Concentration and labor earnings." *American Economic Review* 56:96-117.
- Woodward, Joan
1965 *Industrial Organization: Theory and Practice*. London: Oxford University Press.

CULTURAL EFFECTS ON ORGANIZATIONAL STRUCTURE: THE CASE OF JAPANESE FIRMS IN THE UNITED STATES*

JAMES R. LINCOLN

Indiana University, Bloomington

JON OLSON

University of Southern California

MITSUYO HANADA

Institute of Business Administration and Management, Tokyo

American Sociological Review 1978, Vol. 43 (December):829-847

This paper reports the findings of a survey of 54 Japanese business organizations in the southern California region. We conceptualize these organizations as hybrids whose strategies of structuring work and authority relations incorporate both Western and Japanese patterns. Following Yoshino's treatment of Japanese multinationals, we hypothesize that the extent to which an organization of this sort adopts structural features which are characteristically Japanese depends on the extent to which it recruits Japanese nationals and Japanese-Americans as employees. Based on a series of regressions in which the Aston measures of structure are taken as dependent variables and the proportions of employees who are Japanese and Japanese-American are independent variables along with measures of size, automation, function, and status, we attempt to test this hypothesis. We find relatively strong evidence that functional (i.e., occupational) specialization varies inversely with the size of the Japanese cultural presence. With respect to centralization of decision making, formalization of rules and procedures, and vertical differentiation, only weak and inconclusive tendencies appear in our data.

Introduction

A long-standing proposition of considerable importance to the sociology of formal organizations posits a relationship between culture and organizational structure (Crozier, 1970; Evan, 1974; Hickson et al., 1974). The concept of a cultural effect in essence means that societies vary in the work and authority arrangements of which their formal organizations are composed, and these variations reflect their distinctive traditions, values, and historical experiences. This paper reports on a survey of 54 Japanese business organizations

presently located in southern California. Our primary concern lies in detecting evidence of such a cultural effect in the structural patterns internal to these firms. We believe that multinational organizations in general and these in particular are hybrids of a sort: they adopt in varying combinations the practices of the host society in which they operate, while at the same time retaining some qualities distinctive to their country of origin. We suggest that the extent to which the subsidiary of a Japanese firm conducting business in the U. S. selects its employees from Japanese nationals or Japanese-Americans is an indicator of its cultural bond to Japanese society. By studying the covariations between these indicators and measures of organizational structure, we hope to evaluate certain hypotheses regarding the uniqueness of Japanese organizational forms and their transplanting to the American setting.

Few writers have posed the matter of cultural effects in terms so strong as those used by Crozier (1964:213-27) to describe the character of French bureaucracy. In

* Address all communications to: James R. Lincoln; Department of Sociology; Indiana University; Bloomington, IN 47401.

We are indebted to Koya Azumi and two anonymous ASR reviewers for their constructive comments on a previous draft. This paper was presented to the annual meeting of the American Sociological Association, San Francisco, August 1978. This investigation was supported by Biomedical Sciences Support Grant No. 5-07-RR07010 from the Division of Research Resources, Bureau of Health Professions Education and Manpower Training, National Institute of Health.

his studies of a French clerical agency and a state-owned tobacco monopoly, Crozier sees organizational structure evolving mainly in response to certain features of French national character. Considering three structural properties—centralization of decision making, codification of procedures in formal rules, and the presence of hierarchical barriers to the flow of communications—Crozier contends that the extreme levels he believes are attained by French organizations on these dimensions derive from a culturally determined propensity for Frenchmen to shun face-to-face relationships of personal dependency (1970:57):

Face-to-face dependence relationships are difficult to bear in the French cultural setting because Frenchmen have a very absolutist conception of authority. While they cannot bear omnipotent authority they feel it is indispensable if any kind of cooperative activity is to succeed. "Bureaucratic" patterns of action . . . provide the best solution to this basic dilemma while they tend to perpetuate it.

Notwithstanding Crozier's view of the important role played by culture in the development of French bureaucracy, relatively little attention beyond his own efforts has been paid to the French case. Japan is the target for most research on the problem of culture and complex organization (Abegglen, 1958; Cole, 1971; Dore, 1973; Dowdy, 1973; Nakane, 1970; Rohlen, 1974; Vogel, 1975; Yoshino, 1968). By virtually all accounts, the design and operations of Japanese organizations deviate in important respects from those indigenous to the West. Like Crozier's perspective on the French system, a central theme in treatments of Japanese bureaucracy is the extent to which it embodies patterns which owe their origins to that country's feudal traditions. This view, put forth in Abegglen's (1958) discussion of the Japanese factory and widely held for a time thereafter, has come under growing criticism in recent years. Dore (1973:chap.15), for example, argues that much of what is unusual in the Japanese case is attributable to the phenomenon of late industrialization and is therefore not cultural in the usual sense. Others point to the experience of allied

occupation and the reforms imposed by the United States as developments which left their mark on the structure of the Japanese economy (Yoshino, 1968:29–34). Yet another view is that certain values held over from Japanese feudalism were influential in the creation of a paternalistic employment system and other practices identified as peculiarly Japanese, but this came about as the result of deliberate attempts by Japanese capitalists to legitimate their dominance over workers through appeals to traditional values (Cole, 1971:4; Yoshino, 1968: chap. 3; Kuniyoshi, 1977). They are thus better understood as managerial ideologies in Bendix's (1956) sense than as cultural patterns internalized by individuals in the fashion implied by Crozier.

In this study, our concern is less with the question of whether residues of Japanese feudalism survive in modern Japanese organizations than with the logically prior question of whether Japanese organizations are demonstrably unique. Thus, Crozier sees French organizations as highly centralized, formalized, and stratified merely because they are French and apart from the constraints imposed by size, technology, or function. Many have reached a similar conclusion regarding the form of Japanese organizations: that by virtue of their Japanese origins they present a configuration of structural elements which differs from that typical of organizations native to Europe and North America. It is critical to our approach that we be able to identify that configuration with respect to a set of measurable organizational variables. We therefore will consider four structural dimensions which have been widely studied in contemporary organizational research and for which relatively good measures exist. Using the labels which correspond to their empirical indices, we focus on functional specialization, centralization of decision making, formalization of rules and procedures, and hierarchical or vertical differentiation. We will have occasion, however, to note that in regard to Japanese organizations these terms as conventionally applied are potentially misleading.

1. *Functional specialization.* Weber held that the differentiation of tasks into

formal occupational roles (offices) was a central feature of rational bureaucracy, and few would dispute the view that the occupational division of labor is a highly developed property of Western social organization. But another consensus may be found among students of Japanese social structure: this form of functional specialization is less developed in Japanese organizations (Dore, 1973: chap. 9; Nakane, 1970; Yoshino, 1968: chap. 7). Occupational positions in the sense of sets of functional responsibilities assigned to individuals tend to be ambiguously defined. While strong ties bind the employee to the work group, department, or the enterprise as a whole, assignments to formal occupational roles are unlikely to constitute an important dimension of the division of labor within the organization.¹ Several writers (Dore, 1973:224; Yoshino, 1968:202) have noted that organization charts for Japanese firms are presented, not as networks of individual positions, but as systems of such collective units as divisions, departments, or sections. A thoroughly *cultural* explanation for the low importance attached to occupation as an organizing principle in Japan is the central theme of Nakane's *Japanese Society*. The strong attachments which Westerners have to the concept of position (e.g., occupation, class, status), she calls an *attribute* orientation. This contrasts with the *frame* or collectivity orientation of the Japanese which associates the individual, not with a position or role, but with the entire social group, whether family, community, or company.

2. *Centralization of decision making.* Observers of Japanese management are similarly agreed that a distinctive style of decision making characterizes Japanese formal organizations (Dowdy, 1973; Vogel, 1975; Yoshino, 1968). The *ringi* system is a set of decision-making procedures which combine high centralization of formal authority in top executive posi-

tions with high levels of actual or informal participation by employees of widely varying rank. The distinguishing feature of the *ringi* system is a process whereby proposals originate with lower-echelon employees and are communicated to high-ranking company officials for their approval after a long and complicated process of consensus-seeking discussion and modification in which virtually everyone affected is involved. Although these arrangements have come under attack by Japanese managers and other observers concerned with their cumbersome character, Yoshino (1968:262-72) points out that most such proposals for change call essentially for adjustments in the *ringi* system to improve its efficiency. Few of the system's critics seem prepared to dismantle it altogether.

3. *Vertical differentiation.* If horizontal or task roles are ill-defined in Japanese organizations, the opposite would appear to be true of vertical ranks. Indeed, an elaborate system of status divisions is generally deemed another distinguishing characteristic of Japanese complex organization (Abegglen, 1958:83; Dore, 1973:224; Yoshino, 1968:205). Not only do ranks proliferate, but, as Nakane (1970: chap. 2) demonstrates, direct communications between equally placed organizational units (whether individuals or departments) are minimized. Such exchanges are mediated for the most part through vertical links to superiors. Perhaps the most commonly noted property of Japanese ranking systems is the priority given to age and seniority as promotion criteria over merit and performance. Some observers have suggested that the Japanese tendency to combine diffuse functional responsibilities with fine status gradations derives from, or is at least consistent with, these principles of advancement. They permit senior employees to enjoy superior rank without requiring that they also possess greater technical expertise than their younger subordinates (Nakane, 1970:69).

4. *Formalization.* The extent to which Japanese organizations are governed by formal rules and procedures is not a matter which has been given the same attention received by the foregoing structural

¹ A closely related pattern is the tendency for companies in Japan to recruit new college graduates who have had a general education and train them for careers in the organization. This policy contrasts with one of hiring experienced workers with specialized expertise in particular fields (Yoshino, 1968: chap. 8).

properties. Our expectations concerning this dimension, in fact, are formed more by what we see to be the general drift of writings on the subject of Japanese systems than by a review of literature specifically addressed to the topic of organizational rules. But with respect to a broad conception of formalization, i.e., as formal mechanisms governing behavior in organizational settings, it is not difficult to identify the prevailing point of view. Recall Crozier's assertion that formalization is an alternative strategy of control to face-to-face relations of personal dependency. That belief is echoed in a number of discussions of Japanese organization. But if the French devise formal structures in order to minimize the prevalence of such relations, the Japanese are seen to shun them precisely because dependency ties of this sort are an integral part of Japanese social structure, both within and without organizational boundaries. As Rohlen (1974:118-9) concludes from his ethnography of a Japanese bank:

The acceptance of dependency [with limits] is a definite part of successful personal relationships in Japan, and there is much evidence that dependency is often regarded as a positive aspect of social relations. . . . The patterns [of authority and work relations] we have described are characterized by involvement, an emphasis on feeling, and very little capacity to separate the people from their roles.

Rohlen's views are closely aligned with what is possibly the most general statement of the principles on which Japanese organizations are founded: they count heavily upon a consensus among their members with respect to a set of organizational values and interaction patterns which generate compliance, commitment, and performance without the constraints of official rules and procedures.² In

² That personal, paternalistic authority relations still seem to be preferred by Japanese employees to more structured, impersonal styles of supervision is suggested by Marsh and Mannari's (1976:194) recent survey of three Japanese companies. It was their finding that greater than 80% of the workers in each organization favored a superior who would make demands on subordinates which exceeded the latter's responsibilities under the organization's rules but who would also look after them personally in matters unrelated to work.

Yoshino's (1975:165) view, such consensus is possible because of the homogeneity of Japanese social structure and culture:

As a means for organizing and motivating people, [the Japanese management system] relies more heavily on culturally induced values and less on explicitly defined formal organizational mechanisms. This presents a striking contrast to the American pattern where, among other things, the need to cope with rather heterogeneous and diverse values gave rise to a management system which relies heavily on formal and explicitly defined organizational structures and control systems.

The Structure of Japanese Multinationals

Our interest in this investigation is not with Japanese organizations situated in that country but with their branches and subsidiaries in the United States. However, our expectations concerning the systems adopted by Japanese firms in the United States are heavily conditioned by the portrait of Japanese organization painted above. Attention to the character of Japanese organizations in the American setting is called for if only because of the rate at which their numbers are increasing. In the 1970s, the rate of growth in Japanese overseas investments has outstripped that of any nation: approximately 25% annually. By 1980, only West Germany and the United States are expected to have larger shares of total overseas activity. Today, Britain and the United States are the two largest recipients of Japanese investment, and in financial and commercial industries, the United States alone receives 50 to 75% of all worldwide Japanese investments. To direct attention to the total amount of Japanese business activity in the United States is, however, to greatly underestimate its impact on particular regions. Japanese business is highly concentrated in California which houses the headquarters and branch offices of over half the Japanese firms in the United States. The Los Angeles area, in particular, with two-thirds of the Japanese companies in the state, has felt the expansion of the Japanese presence.³

³ These figures are taken from a privately commis-

Given the rising visibility of Japanese firms on the American landscape, the question of how they structure work relations takes on new importance. Such enterprises are employing United States citizens in growing numbers. How the foreign corporation and its local labor force react to one another may depend in large measure on the degree to which genuine cultural differences exist between the two countries with regard to the expectations employees hold concerning organizational life. Although one recent study (Johnson and Ouchi, 1974) concludes that job satisfaction among the American employees of Japanese owned and managed companies is high, reports of cross-nationality frictions among the employees of United States-based Japanese firms point to a less positive conclusion.⁴

Certainly many distinctive features of the Japanese employment system are not imported to the United States when a Japanese firm sets up operations and hires employees here. Locally-hired workers are not guaranteed lifetime employment nor do they share in the wide-ranging fringe benefit and company-sponsored recreational and social programs so often described in the literature on employment relations in Japan. But as far as structure is concerned, there are indications that certain typically Japanese approaches to organizing work relations find their way into the subsidiaries of Japanese companies which operate in the United States. Two authors of recent detailed studies of Japanese multinationals believe that traditional Japanese practices survive in the foreign subsidiaries (Tsurumi, 1976:102; Yoshino, 1976:167). Tsurumi further argues that their presence in the United States poses a serious challenge to well-entrenched American business ideas concerning efficient management. The re-

search conducted by Johnson and Ouchi (1974) also turned up traces of Japanese practices in the United States subsidiaries of Japanese firms. These included an emphasis on consensus and participation in the decision-making process and a tendency to meet new task requirements by shifting employees from one job position to another instead of taking on new personnel.

Of particular importance for the strategy we take in this investigation is Yoshino's (1975:164-5; 1976:166-78) observation that the perpetuation of Japanese organizational structures in the overseas branches of Japanese firms depends heavily on the presence of Japanese nationals in the foreign branch.⁵ In his view, Japanese managerial practices and organizational forms are not consciously implemented devices but rather reflect the socialization undergone by Japanese employees in the organizational culture peculiar to that society. Thus, Japanese managers heading overseas operations are predisposed toward instituting those procedures with which they are familiar and eschewing those with which their experience is limited. Furthermore, Japanese organizational arrangements require the maintenance of informal, personal networks both among the employees of the foreign branch and between them and the management of the parent firm in Japan. Non-Japanese employees would find such ties very hard to acquire.

In California, where a relatively large

sioned report on the extent and growth of Japanese investments in the United States as a whole and in California in particular. See also Tsurumi (1976).

⁴ For example, the *Los Angeles Times* (February 19, 1975) reported on a court suit brought by American employees against a Texas-based Japanese firm. Among other things, the suit alleged that Americans were "excluded from important night meetings held by the Japanese, because the American wives 'would not understand or condone such evening work,' an old Japanese tradition."

⁵ The author of a recent study comparing the communications and decision-making practices of managers in American companies and Japanese companies operating in Japan and the United States concludes quite differently that the proportion of Japanese nationals in the U.S. subsidiary of a Japanese firm is not, on the whole, related to managerial behavior (Pascale, 1978:108). While these findings are not what we would have expected, they do not speak directly to the issues we address since our measures pertain to global properties of structure, not orientations of individual managers. (Pascale does present an appendix describing certain structural differences between American and Japanese firms in the United States but does not discuss them in the text.) Contrary to the perceptions of writers like Tsurumi and Yoshino, Pascale (1978:105) finds that, with respect to communication behavior, Japanese firms in Japan differ more from either Japanese or American firms in the United States than the latter do from each other.

segment of the population is of Japanese descent, Japanese subsidiaries have shown a strong tendency to hire Japanese-Americans. The possibility exists, then, that the cultural orientations which Yoshino suggests contribute to the administration of the organization along Japanese lines may be found among this employee group as well as that comprised of Japanese nationals. In fact, we are inclined to believe that some subsidiaries view the hiring of Japanese-Americans as a kind of compromise between the pressures to take on local labor while preserving the Japanese character of the firm.

Whether Japanese-Americans are carriers of Japanese culture is, of course, a proposition open to question. Most observers agree that Japanese-Americans have adapted well to life in the United States but disagree on whether the reasons lie in a "good fit" between Japanese culture and American society or in a willingness by Japanese-Americans to give up their traditions in favor of United States middle-class values (Caudill and DeVos, 1956; Connor, 1977; Maykovich, 1972). Kitano (1969:117-9) argues that the rapid social change experienced by twentieth century Japan has made the values of traditional Japan more prevalent among Japanese-Americans than among the Japanese themselves. He identifies the minority status and relative isolation of the Japanese-American community as conditions which have preserved habits and customs no longer prevalent in Japan.

The accuracy of Kitano's portrayal of Japanese-Americans in general may be debatable, but if true at all one would expect it to best describe the inhabitants of such relatively segregated Japanese-American communities as Little Tokyo in Los Angeles. It is precisely from such ethnic enclaves that the United States branches of Japanese firms tend to recruit Japanese-Americans as employees. Japanese-Americans are an important consumer market for Japanese firms in southern California, the banks especially being concentrated in Japanese-American neighborhoods. As Table 2 shows, a higher fraction of bank personnel are Japanese-American than is true of other enterprise types.

We believe that our approach—that of measuring the changes in organizational structure associated with rising tendencies to employ Japanese nationals and Japanese-Americans—is a fairly direct way of assessing the impact, if any, that Japanese culture has on organizational structure. If individuals are carriers of culture, a shift in the direction of more Japanese and fewer non-Japanese should imply stronger pressure to fashion the workings of the organization along Japanese lines. We also should reiterate our conviction that these are hybrid organizations: neither Japanese nor American entirely but falling between these ideal types. Position on this continuum, then, is indexed by the national origins composition of employees. That conviction was strongly reinforced by our visits to the sampled firms. We cannot describe adequately how different is the atmosphere in an organization where 50 to 80% of the personnel have Japanese origins, where Japanese is widely spoken and certain Japanese interpersonal customs are observed, from one employing few or no Japanese or Japanese-Americans. Both are Japanese subsidiaries but one is profoundly more "Japanese" than the other. Thus, we have two perspectives on these issues. While we do not rule out the possibility that the presence of Japanese employees per se is causally linked to the structure of the firm, we assign equal if not greater importance to a view which sees this variable as essentially an indicator of a much more general cultural dimension along which the organizations in our sample may be arrayed.

Research Design

The findings reported here were produced by an interview survey of Japanese business organizations in southern California. Our sample is composed of 54 organizations which are heterogeneous in function (i.e., industry) and organizational status (whether the unit is a subsidiary, branch, etc.). The sample was purposively drawn from the roughly 430 Japanese firms in the region with the intent of selecting reasonably equal proportions from major industry groups and a representa-

tive distribution with respect to size. Nearly 85% of the organizations contacted produced usable interview data. The interviews were conducted with chief executives and/or their appointed assistants and, with the exceptions of those few cases where such persons were not Japanese nationals, in Japanese. The capacities of our informants to converse in English typically were exceeded by the complex questions posed in the interviews.

The instruments we used to measure structural properties are based on the Aston scales (Pugh et al., 1968; 1969). Although developed for use in Britain, these measures have been applied to sets of organizations in several national contexts. We employed the short forms of the Aston scales which Inkson and his colleagues (1970) have shown to be reliable substitutes for the original longer scales.

It is important that we comment on the shortcomings of an approach which adapts a set of measures designed for use in one country to cross-cultural research. Not only do the Aston instruments reflect only formal dimensions of structure (i.e., those revealed in official documents or the perceptions of managers), but the constructs they represent are narrowly conceived; more so than their broad labels of centralization, formalization, and so on, would suggest. Consider the structural differences that are presumed to exist between Japanese and Western organizations. As defined by the Aston scale, functional specialization means the assignment of specialized tasks to individuals. An organization which allocates functions to some other unit would reveal low scores on the Aston index measuring this variable. As seen by many observers, however, Japanese organizations display precisely this pattern of task allocation. The view with which we are familiar is not that Japanese organizations are functionally homogeneous; only that the parts to which tasks are assigned are not offices or occupations and the individuals who fill them. Work groups, departments, and other collective units may be responsible for quite specialized roles.

Similarly, based on our understanding of the ringi system, we would predict that

the Aston centralization measure would show Japanese organizations to have highly concentrated power structures. That scale is designed to detect the average level to which formal responsibility for decisions on a variety of issues is delegated.⁶ Japanese organizations are perceived to delegate little formal responsibility for decisions. Thus, official authority with respect to most questions of policy is monopolized by high-ranking managerial elites. But the essence of the ringi system, as Yoshino (1968) stresses, is the importance it attaches to participation in the fashioning of the decision by all affected personnel. The appearance of a highly centralized decision-making arrangement would thus seem to mask an unusually participatory process.

This is not to say that useful information cannot flow from the application of the Aston scales to organizations whose internal arrangements may incorporate principles which differ from those on which the scales' construction is premised. It is important to know whether Japanese organizations differentiate occupations and delegate decision-making authority to a lesser degree than do Western organizations. But one must be cautious in applying such diffuse terms as centralization and functional specialization to the variables measured on such scales, since these convey meanings which are broader than those which the scales are designed to reflect.

Our study is not, in fact, the first to apply the Aston scales to a sample of Japanese organizations. Azumi and McMillan (1974) surveyed 50 factories in Japan in a recent investigation, producing scores on the scales which correspond to the four structural dimensions discussed earlier in this report. Moreover, in one paper, these writers contrasted means on these variables calculated from the

⁶ Child (1972:174), too, points out that the Aston centralization scale measures delegation of authority and correctly notes that such delegation is compatible with Weber's model of bureaucracy, not contrary to it as argued by the Aston group (Pugh et al., 1968:82). In Weber's treatment, bureaucracies centralize power in the sense that the highest source of authority (that to which decisions at lower levels can be appealed) is vested at the apex of a pyramid of offices.

Japanese sample with similar values produced in two British studies: the Aston survey of 46 organizations (Pugh et al., 1968) and the National sample of 82 units (Child, 1972). Azumi and McMillan report that the Japanese and British samples differ on the same measures so as to indicate that the Japanese organizations are less specialized, more centralized, more vertically differentiated, and more formalized than the British organizations.

With the exception of the finding pertaining to formalization, these are the results we would have anticipated.⁷ Azumi (1978) reports, however, that a subsequent comparison of matched (for size and product) samples of 12 organizations each taken from Japan, Britain, and Sweden revealed little between-country variation on all structural dimensions except the number of ranks. In that case, the Japanese organizations again proved to be more highly differentiated. Such comparisons are, of course, risky given the small number of cases and the possibility that cross-national differences may yet be confounded with functional or technical characteristics of the units studied. But regardless of what cross-national research reveals about the variation (or its absence) in organizational structure from one country to the next, we believe that the approach we have chosen—to inquire whether structural differences among a set of multinational organizations depend in any systematic way on the extent of social and cultural integration of such units in the host society—is a valuable strategy in its own right and a useful supplement to cross-national research.

The specific variables which we will analyze here and whose values were produced through the interviews with key executives are as follows: In addition to the information conveyed to us by informants, organization charts, company handbooks, and other documents describing the firm or branch also were collected and used to cross-validate the interview data.

⁷ Dore (1973:243) similarly reports from his comparison of two Japanese electronics factories with two similar British organizations that the Japanese company formalized more procedures in written rules and regulations than did its British counterpart.

1. **SIZE**: the total number of employees in the unit studied. As usual in organizational research, **SIZE** is quite skewed toward high values in this sample and tends to be nonlinearly related to the measures of organizational structure. The regression and correlation results to be reported here are therefore based on a logarithmic transformation of this measure.

2. **AUTO**: the *automaticity range* index developed by the Aston group to measure technological complexity. It indicates the most automated piece of equipment in use by the organization on an ordinal scale of one to five.⁸

3. **JAPAN**: the proportion of the organization's employees who are Japanese citizens.

4. **JAMER**: the proportion of the organization's employees who are American citizens of Japanese descent.⁹

⁸ The Aston group employed two measures of automation: *automaticity mode* (the automation score of the bulk of equipment in use by the organization) and *automaticity range* (the highest scoring piece of equipment). In this sample, however, 42 or 79% of the 54 organizations scored two (see below) on automaticity mode. This distribution reflects the fact that manufacturing firms constitute a small proportion of the sample, and the nonmanufacturing organizations are heavily engaged in information-processing activities (e.g., billing and accounting). Among the latter, electric typewriters and desk calculators constitute the bulk of the machinery, while computers are the most automated machines used. The levels of this dimension and its coded scores are:

Hand tools and manual machines	1;
Power machines and tools	2;
Single-cycle automatics and self-feeding machines	3;
Automatic: repeats cycle	4;
Computer control: automatic cognition	5.

This scale is not identical to the original Aston measure (see Inkson et al., 1970), for we have omitted one category (self-measuring and adjusting: feedback) which held no cases.

⁹ The criterion we used to distinguish between Japanese nationals and Japanese-Americans was simply citizenship. Our informants could not give us reliable reports of the generation of the Japanese-Americans; thus, we cannot partition this category into *Isei*, *Nisei*, and *Sansei*. We were, however, able to divide the Japanese nationals into those who held a permanent U.S. visa and those who did not; a distinction which implies some difference as to time spent in and assimilation to the United States. Treating these subgroups of **JAPAN** as a separate variable made little difference in the results we obtained. Nor

5. SPECIAL: the 16-item Aston scale designed to measure the extent to which organizational functions are assigned as specialized duties to individuals. Cronbach's alpha computed for these items yields a reliability value of .835.

6. CENTRAL: a 23-item scale designed to measure the level in the official hierarchy of authority given final responsibility for decisions. Alpha = .886.¹⁰

7. FORMAL: the 12-item *formalization of role definition* scale devised by the Aston group. The measure is based on the number of specific role-defining documents in the organization and the degree to which these are distributed among employees. Alpha = .739.

8. RANKS: the *vertical span* measure devised by the Aston group. It is the number of job positions in the longest line

did combining JAPAN and JAMER in a single measure of "Japaneseness" alter the conclusions indicated by our data. We also recognize that the use of ratio variables such as JAPAN and JAMER in regression analysis has been the object of some criticism in the sociological literature (e.g., Fuguitt and Lieberman, 1974; Schuessler, 1974). While it is in fact the *ratio*—the number of Japanese employees relative to the organization's total labor force—which is the relevant theoretical variable here not the number of Japanese or Japanese-Americans per se, we did estimate the same regressions substituting the absolute values of these variables for their corresponding proportions. This transformation had the effect of heightening the contrast between the regression for SPECIAL and those for the remaining structural variables. Strong, statistically significant negative effects materialized in the SPECIAL equation, whereas the coefficients for these regressors in the other three equations differed only trivially from zero.

¹⁰ Our centralization scale contains the same items as the *autonomy* scale recommended as an abbreviated measure of centralization by Inkson et al. (1970), but it is scored differently. Their index consists of 23 items representing decisions which are coded dichotomously as made within or outside the organization studied. We coded each decision according to the level in the organizational hierarchy at which it was made, as follows. Essentially the same coding is used in the longer version of the Aston centralization scale.

- 6 Made in higher organizational unit;
- 5 Head of division or above;
- 4 Department head;
- 3 Assistant manager;
- 2 Officer, specialist, professional, supervisor (white-collar);
- 1 Foreman, supervisor (blue-collar);
- 0 Direct workers.

Table 1. Joint Frequency Distribution of Sampled Organizations by Status and Function

Function	Status			Total
	Agencies	Subsidiaries	Branches	
Banks	2	4	10	16
Trading	0	0	8	8
Manufacturing	1	12	1	14
Distribution	1	11	4	16
Total	4	27	23	54

$\chi^2 (6) = 24.851$.

between the lowest level employee and the top-ranking executive.

Analysis and Findings

Table 1 indicates how the 54 organizations in this sample are distributed by status and function. We estimate that the four functional categories in which these units are grouped account for greater than 75% of the Japanese firms in the Los Angeles region. The remaining 25% consists primarily of small retail shops and service establishments. Of this 75%, we estimate that in the region as a whole, their industry distribution is approximately as follows: banks (15.8%), trading companies (17.9%), manufacturing (13.1%), and distribution (53.1%). Comparing these figures to those in Table 1, one can see that banks and manufacturing organizations are overrepresented in the sample while trading firms and distributors are correspondingly underrepresented.¹¹

As we use the term, *status* denotes a three-way breakdown of organizational units. *Agencies* are branches of the Japanese parent company which are not incorporated in the United States. *Subsidiaries* are legally incorporated subsidiaries of the Japanese company.

¹¹ The data pertaining to the population of Japanese organizations in southern California were made available to us by the Japan Trader's Club of Southern California. The figure of 430 units is based on all addresses of Japanese-owned establishments.

Table 2. Raw and Adjusted Means on Organizational Variables by Categories of Status and Function †

Variable	Status			Function					Total	
	Agency	Sub-sidiary	Branch	η (β)	Banking	Trading	Manu-facturing	Distri-bution	η (β)	SD
SIZE	24.00 (33.45)	131.40 (117.03)	58.22 (73.50)	.43 (.35)	34.19 (57.11)	56.50 (75.30)	99.29 (83.64)	162.19 (153.57)	.33 (.22)	92.30 121.58
AUTO	3.50 (3.55)	4.37 (4.54)	4.52 (4.32)	.31 (.31)	4.37 (4.47)	4.75 (4.80)	3.86 (3.78)	4.62 (4.58)	.40 (.44)	0.83
JAPAN	0.55 (0.56)	0.18 (0.24)	0.25 (0.18)	.51 (.51)	0.31 (0.30)	0.35 (0.41)	0.17 (0.15)	0.30 (0.30)	.40 (.50)	0.19
JAMER	0.03 (0.01)	0.08* (0.11)	0.27 (0.24)	.55 (.41)	0.29 (0.27)	0.23 (0.15)	0.06 (0.10)	0.08* (0.10)	.57 (.42)	0.18*
SPECIAL	1.50 (1.59)	5.50 (5.35)	1.52 (1.68)	.60 (.56)	2.13 (3.05)	1.38 (3.81)	4.43 (3.90)	5.16 (4.47)	.46 (.19)	3.51
CENTRAL	116.25 (113.52)	111.63 (111.82)	119.02 (119.27)	.37 (.37)	123.65 (122.09)	111.43 (107.28)	112.65 (115.30)	110.59 (111.92)	.53 (.53)	115.12 9.74
FORMAL	11.75 (11.73)	11.08 (11.71)	9.77 (9.03)	.18 (.34)	10.93 (11.47)	10.21 (11.76)	9.71 (8.76)	11.13 (10.66)	.15 (.29)	10.57 3.92
RANKS	4.00 (3.93)	5.67 (5.78)	4.70 (4.57)	.35 (.42)	5.13 (5.46)	4.63 (5.18)	4.36 (4.92)	5.19 (4.95)	.14 (.14)	5.13 1.61
N	4	27	23		16	8	14	16		54

† Adjusted means are in parentheses.

* These values are based on one less than the number of cases indicated because of a missing observation on JAMER.

Branches are branches of the United States subsidiary. The majority of these units is almost evenly divided between subsidiaries and branches. Only four agencies were among the organizations surveyed.¹²

It is clear from Table 1 that function and status are hardly independent dimensions. Although most of the banks are branches, they are more evenly distributed across levels of status than are other industry types. With only two exceptions, the manufacturing firms are full subsidiaries, as are most distribution companies. In order to separate the effects of these two classifications on the variables we have measured, Table 2 presents both unadjusted and adjusted means on these variables for each category of status and function. The adjustment produces a hypothetical mean which one would expect to observe were the classification dimensions independent. Presenting our results in this fashion is considerably simpler than a full cross-classification of means on all eight variables and conveys the same information since tests for statistical interaction in the status and function effects proved negative without exception.

We should note at the outset that these organizations are considerably smaller than those in other samples to which the Aston measures have been applied. We explored the possibility that the presence of very small units might produce unusual patterns among these variables by deleting the bottom quartile of the size distribution (range: 7–24 employees) and reanalyzing the data. These results, however, did not differ materially from those based on the total N of 54. The size distribution of these firms also reveals some definite variations by status and function. Banks, averaging

34 employees, are smallest, followed in order of size by trading, manufacturing, and distribution. Also, as might be anticipated, subsidiaries are considerably larger than branches or agencies.

It is further important that we examine the employment composition of these firms as regards Japanese citizenship and descent. The mean ratio of Japanese citizens to total employees in the 54 organizations is .24, while the Japanese-Americans average 16% of employment. This second figure pertains to 53 units. One organization was not able to give us a reliable count of the number of Japanese-Americans it employed. Not surprisingly, agencies of the parent company in Japan employ the largest proportions of Japanese nationals, whereas Japanese-Americans are best represented in branches. With respect to function, Japanese-Americans are most likely to be employed by banks, as we observed previously, while the mean proportion of Japanese-employees is roughly constant at slightly over 30% in banks, trading companies, and distributors. It is about half this figure, however, in the manufacturing firms.

An unanticipated finding in Table 2 is the low mean revealed by manufacturing organizations on the automation variable. Working with the composite index, *work-flow integration*, of which the present automation scale was a component, Hickson et al. (1969) found just the opposite in a sample of 46 British organizations. Manufacturing units were distinguished from their nonmanufacturing counterparts by high scores on this measure of technological complexity. In our sample, the reason for the low mean among the manufacturing units is not hard to discern. Thirty of the 54 organizations scored "5," the highest value possible, indicating that a computer was the most automated piece of equipment in use. Of these 30, only three were in manufacturing. The nonmanufacturing organizations in our sample have heavy data processing requirements, record keeping, accounting, and billing being major parts of their operations. Consequently, the probability that they use a computer is high. Given this distribution on our automation vari-

¹² In the case of three companies, multiple units of the same firm were sampled. Two banking corporations yielded interviews in three offices apiece, while two branches of the same distribution company were included in the sample. Since we might expect structural similarities among the component organizations of a single corporation, we tested for such effects by adding three dummy variables representing those companies to the regressions shown in Table 4. In no equation did these produce effects which were significant at the .05 level, nor did their inclusion alter the coefficients estimated for the other predictor variables.

able, it is essentially the same as the simple dichotomous measure (presence or absence of computer) used by Blau and Schoenherr (1971:127) to tap the same dimension.

A fairly complete picture of how the dimensions of organizational structure vary with status and function can be gleaned from Tables 2 and 4.¹³ We will have more to say about Table 4 later, but for now consider the information it conveys regarding the magnitude and direction of status and function effects net of each other and those of size, automation, and the proportions Japanese and Japanese-Americans. Table 4 presents regression equations for each of the four structural measures. The relevant information appears in equation 2 which, for each dependent variable, includes size, automation, and the two "Japaneseness" indicators as regressors, plus three dummy variables representing function and two representing status.¹⁴ The omitted categories are banking and branches, respectively.

How would we expect organizational structure to depend on status and function? Relatively little systematic attention to these questions can be found in the organizational literature. Organizational researchers in the past have either sampled units which differed on these dimensions then ignored their implications for the findings (e.g., Pugh et al., 1969) or selected homogeneous units thus limiting the generality of their conclusions to organizations of the type studied (e.g., Blau and Schoenherr, 1971). We contend

that the best approach to these matters is that adopted here: (1) to sample organizations which are heterogeneous with respect to status and function; (2) then to assess the variations these factors produce in organizational properties; (3) finally, to estimate the relations among such properties net of status and function effects.

Indeed, so little attention has been paid to status and particularly function effects in the comparative research literature on organizations that appropriate hypotheses are not easy to come by. Child (1972) brought the issue of status to the awareness of organizational analysts in trying to make sense of his finding that the organizations in his "National" sample of British firms displayed a mean on the Aston centralization scale which was notably lower than that obtained in the original Aston sample (Pugh et al., 1968). Child reasoned that the higher proportion of branches in the Aston sample may have produced this discrepancy. For branches, some decisions affecting the organization may be made by the larger unit of which it is a part. Such decisions are coded on the Aston scale as more "centralized" than decisions made by the director of the branch itself. Although some debate has ensued over whether this difference in status composition in fact produced the discrepancy in centralization scores observed between the National and Aston samples (Donaldson et al., 1975) it is clear that status in our sample has precisely the effect diagnosed by Child. Table 2 shows branches to be more centralized than agencies or subsidiaries, and Table 4, which introduces additional controls, reveals the same pattern. The unstandardized regression coefficients in Table 4 which are estimated for AGENCY and SUBSID are interpretable as adjusted deviations of the means of these status categories from that of branches. The negative coefficients paired with these dummy variables indicate that the adjusted means of agencies and subsidiaries on the centralization measure are lower than that of the omitted category, branches. In addition, Table 4 indicates that the branches are less specialized, less formalized, and have fewer ranks than

¹³ We also performed tests of the hypothesis that status and function interact with size, automation, and the measures of a Japanese presence. While the small N's on which the within-group regressions were based preclude strong inferences, we found little overall evidence that an additive model is inappropriate. There did, however, appear to be some tendency for the proportions Japanese and Japanese-American to have stronger effects on the centralization of banks compared with other industries.

¹⁴ Two additional variables we considered in the regressions were the age of the company in the U.S. and the age of the particular unit studied. Neither displayed significant effects on any of the four structural measures, however, and we therefore do not report these results.

Table 3. Zero-Order Correlation Coefficients

	2	3	4	5	6	7	8
1 SIZE	.354	-.610	-.176	.687	-.262	.156	.618
2 AUTO		-.250	.024	.234	.119	.286	.428
3 JAPAN			-.121	-.509	.146	.037	-.398
4 JAMER				-.399	.396	-.249	-.159
5 SPECIAL					-.264	.230	.590
6 CENTRAL						.107	-.128
7 FORMAL							.303
8 RANKS							

agencies or subsidiaries when factors like size and automation are controlled. They are, in other words, generally less bureaucratic in structure.

As for function effects on structure, our expectations are only impressionistically formed.¹⁵ There is a tendency to regard manufacturing organizations as incorporating routine work rhythms and highly structured patterns of supervision, thus rendering them more bureaucratic than organizations in other industries. This impression does not find much support in our data, however. In fact, a clear pattern of variation among industries with respect to structural traits is difficult to identify. Table 2 suggests that banks are more centralized and less specialized than other types, while manufacturing has the lowest score on formalization. When the additional controls in Table 4 are imposed, the banks appear only slightly more centralized than manufacturing firms and distributors. Moreover, in a complete reversal from Table 2, the banks now display the highest adjusted mean on specialization, and they appear to be more formalized and stratified as well. Although these results indicate that industrial differences are not trivial and must therefore be taken into account, we confess to finding no clear pattern of structural variation by function which is not attributable to other sources.

¹⁵ Hickson et al. (1969) and Aldrich (1972) have discussed the implications for the Aston findings which arise from that sample of 46 organizations being divided between manufacturing and service industries. These analysts observe that the correlations between the Aston technology measure, *work-flow integration*, and measures of structure all but vanish when this industrial classification is controlled.

In Table 3, we present the zero-order correlations among the measures yielded from the interviews. We should pause at this point to contrast this pattern of relationships with those obtained by other investigators working with similar measures but different samples of organizations. In general, our results are consistent with those generated by previous studies. Specialization, formalization, and vertical differentiation (RANKS) are positively intercorrelated, while centralization is inversely associated with specialization and vertical differentiation. However, the sign of the correlation coefficient linking formalization and centralization in this sample is opposite to that found in two British samples (see Child, 1972). We would not go so far as to make it the basis for a broad cultural argument, but we do note here with interest that a similar weak positive correlation between this pair of variables showed up in Azumi's (1978) analysis of the Aston measures applied to a sample of 50 factories in Japan.

The findings of earlier investigations also would have led us to anticipate the positive correlations of automation, specialization, formalization and ranks with size. The likewise positive correlations linking specialization, formalization, and number of ranks to automation are again consistent with the findings of the British and other studies. The weak positive association between this measure of technology and the Aston centralization scale is at odds with the inversely signed coefficients for this pair of variables calculated by Pugh et al. (1969) and Blau and Schoenherr (1971:117), but is matched in the findings reported by Child (1972) from his replication of the Aston survey.

Table 4. Regressions of Structural Variables on Proportions Japanese and Japanese-American and Other Organizational Properties

Dependent Variables	Intercept	Independent Variables										R ²	R ² †
		Unstandardized Regression Coefficients											
		SIZE	AUTO	JAPAN	JAMER	TRADE	MANU	DISTRIB	AGENCY	SUBSID			
SPECIAL	1	-0.838 (0.396)	0.030 (0.384)	-4.483 (2.076)	-6.526 (1.734)618	.595	
	2	0.767	-0.208 (0.371)	-5.763 (2.429)	-6.710 (2.181)	-0.931 (0.961)	-3.038 (1.107)	-1.812 (0.914)	1.374 (1.325)	2.629 (0.777)	.736	.688	
CENTRAL	1	107.815	-1.982 (1.655)	2.458 (1.606)	18.405 (7.252)219	.171	
	2	102.549	0.324 (1.516)	3.511 (1.659)	9.429 (8.923)	-16.541 (3.931)	-0.532 (4.529)	-7.291 (3.739)	-6.269 (5.419)	-8.010 (3.178)	.483	.389	
FORMAL	1	2.842	0.413 (0.695)	1.378 (0.674)	-4.271 (3.043)150	.098	
	2	5.638	0.461 (0.751)	-0.099 (4.918)	-6.303 (4.417)	-1.155 (1.946)	-3.317 (2.242)	-2.430 (1.851)	2.435 (2.628)	1.497 (1.573)	.217	.075	
RANKS	1	0.165	0.784 (0.232)	-0.361 (1.217)	-0.692 (1.017)442	.407	
	2	0.369	0.919 (0.228)	-0.540 (1.493)	-1.646 (1.341)	-1.133 (1.341)	-1.082 (0.680)	-1.715 (0.562)	0.162 (0.814)	0.580 (0.478)	.576	.499	

Table 4.—Continued

Dependent Variables	Intercept	Independent Variables							
		Standardized Regression Coefficients (Values Corrected for Attenuation)							
		SIZE	AUTO	JAPAN	JAMER	TRADE	MANU	DISTRIB	AGENCY
SPECIAL	1	.496*** (.541)	.008 (.010)	-.234* (-.279)	-.355*** (-.389)
	2	.482*** (.521)	.052 (-.059)	-.326* (-.368)	-.365** (-.411)	-.100 (-.106)	-.404* (-.456)	-.251 (-.284)	.109 (.122)
CENTRAL	1	-.209 (-.212)	.210 (.222)	.115 (.138)	.342* (.366)
	2	.034 (.051)	.300* (.325)	.290 (.345)	.175 (.205)	-.609*** (-.657)	-.204 (-.003)	-.345 (-.353)	-.170 (-.193)
FORMAL	1	.108 (.125)	.293* (.342)	.151 (.175)	-.197 (-.231)
	2	.121 (.136)	.241 (.281)	-.005 (-.011)	-.291 (-.345)	-.106 (-.121)	-.375 (-.445)	-.286 (-.339)	.164 (.194)
RANKS	1	.499**243*	-.042	-.078
	2	.586***256*	-.063	-.185	-.272	-.297	-.489*026

† \bar{R}^2 is R^2 adjusted for degrees of freedom.* $p < .05$.** $p < .01$.*** $p < .001$.

We believe that this descriptive discussion of our data was in order so that we might locate our results with reference to the existing research literature on organizations and address what we consider to be the important question of status and function variations. We focus now on the key problem of the investigation: to ascertain whether the structural properties of these organizations vary with the proportions of Japanese nationals and Japanese-Americans they employ in a fashion consistent with our earlier arguments. Returning to Table 4, let us examine the effects of variables other than the status-function dummies. Besides the fairly conservative (given 54 cases) decision rule of statistical significance for attributing importance to effects, we may wish to consider the standardized regression coefficients, both raw and corrected for measurement error attenuation, which are presented in the Table.¹⁶

Taking first the case of SPECIAL, we observe that equation 1 reveals the strong positive effect of size so often reported in comparative organizational research. Automation, however, has no net influence on this structural variable. But the effects of our two indicators of Japanese presence are quite pronounced. The data strongly support the hypothesis that the degree to which individuals are assigned functionally specialized roles in the organization will diminish as the proportions of Japanese and Japanese-Americans rise. Both JAPAN and JAMER have statistically significant negative effects on the functional specialization scale. Equation 2 demonstrates that the coefficients estimated in equation 1 are not biased by the omission of controls for function and status. The effect of size is still pronounced, that of AUTO is negligible, while the measures of "Japaneseness" retain significant negative coefficients. Indeed, whether standardized or unstan-

dardized coefficients are compared, one is struck by the similarity in the calculated effects of JAPAN and JAMER.

Equation 1 shows CENTRAL to be positively dependent on automation and negatively on size, but neither of these effects are significant at the conventional 5% level. The coefficients paired with JAPAN and JAMER have the expected signs. Both are positive: the pattern we anticipated finding were decision making in firms employing high proportions of Japanese and Japanese-Americans more likely to approximate the ringi model. Yet only in the case of JAMER is this effect significant. When we shift attention to equation 2, however, the picture changes. While neither coefficient exceeds twice its standard error, the larger (standardized or unstandardized) coefficient is that of JAPAN.

The evidence for our last two hypotheses—that predominantly Japanese organizations should prove to be less formalized but more vertically differentiated—is thin to negative. None of the regression coefficients in the equations for FORMAL is significant, and, while its standardized regression coefficient is large enough that we would not dismiss the negative effect of JAMER, that of JAPAN has the wrong sign in equation 1 and goes to zero in equation 2. As for RANKS, our hypothesis is flatly contradicted: the effects of both indicators of Japanese cultural presence are *negative*.

One might argue that the reasons to expect rank differentiation in the U.S. subsidiaries of Japanese firms to rise as they vary from less to more Japanese are not particularly strong. If, as we suggested earlier, status differentiation in Japanese organizations reflects the operation of the seniority principle in separating persons by rank, the fact that seniority in the subsidiary does not appear to outweigh merit or talent as a criterion for pay or promotion would obviate the grounds for our hypothesis. With the exception of Japanese nationals transferred by the parent firm (who remain classified under its seniority system), there is little evidence that the U.S. subsidiaries of Japanese corporations differ much from firms native to this country in pegging status

¹⁶ Based on the reliability estimates for the summed scales (SPECIAL, CENTRAL, FORMAL), correlation coefficients involving these variables were corrected for attenuation using the standard formula (see Bohrnstedt, 1970:84). The corrected standardized regression coefficients reported in Table 4 were calculated from these adjusted correlations.

mainly to training and merit. Even among the Japanese transfers, the seniority classification of the parent firm is a latent hierarchy which need not be isomorphic with the subsidiary's vertical division of labor. Perhaps the number of ranks in this second hierarchy rises with the proportion of Japanese citizens in the U.S. firm, but we do not have the data available to test that hypothesis.

CONCLUSIONS

Our strategy in this research was to conceptualize Japanese firms conducting business in the United States as organizations placed on a cultural continuum from more to less Japanese. Location on that continuum was indexed by the extent to which such firms employed Japanese nationals or Japanese-Americans instead of employees with no Japanese background. This approach is consistent with Yoshino's argument that the presence of Japanese nationals is reflected in the structure of the foreign branch and Kitano's view that Japanese cultural traditions survive among Japanese-Americans. We should note that an indirect test of the assumption that the presence of Japanese-Americans reflects a Japanese cultural influence in the firm is afforded by a comparison of this variable's effects with those of JAPAN. We conclude that the outcome of that test is positive: in general, JAPAN and JAMER behave very similarly with respect to the structural measures.

Firm evidence in our data for the existence of a cultural effect, however, is present only in the case of functional specialization. As a wide-ranging literature on Japanese social structure led us to expect, the degree to which narrowly-defined occupational roles are found in these transnational organizations varies inversely with the size of the Japanese cultural presence. No significant effects of tendencies to employ Japanese and Japanese-Americans could be isolated with respect to centralization, formalization, and vertical differentiation, although the coefficients calculated for the proportions Japanese and Japanese-American in

the equation predicting centralization had the signs we anticipated.

Let us briefly discuss some implications of these findings for organizational theory. Some students of organization (Hickson et al., forthcoming) argue that the critical question in cross-cultural research is whether organizational theories proved valid in one national setting are equally applicable to others. Should the interrelations among dimensions of structure or their common dependencies upon such postulated determinants as size or technology prove to fluctuate from one society to the next, the goal of developing general organizational theory would appear to be unattainable. In fact, findings from a rapidly growing number of replications of organizational surveys using the Aston scales and other measurement devices so far point to little evidence for the existence of such nonadditive cultural effects (Hickson et al., 1974; McMillan et al., 1973; Tracy and Azumi, 1976).

But the question of whether culture makes a significant net contribution to variation in organizational arrangements *in addition* to those produced by technical or functional constraints is hardly a trivial matter, for it points to the conclusion that rationalistic theories of organizational action and adaptation are at best incomplete. Since the focus of most comparative organization research is on dimensions of *formal* structure, a recognition that organizations incorporate as permanent features of their operations blueprints the habits, customs, and values of their members and the larger populations in which they are embedded seems particularly in order, a point well developed and explored by Meyer and Rowan (1977). Given the considerable slippage which organizations can apparently tolerate between those official designs and the actual interaction structures which they develop in the course of day-to-day activities, it may be in the organization's best interests to make its public face conform as best it can to prevailing ideological or normative currents and cope in an ad hoc fashion with the strains that such slippage creates.

While we intend these observations as a call for increased attention to the processes whereby organizations assimilate

elements of their sociocultural environments, we find that they are also compatible with the proposition that Japanese organizational styles may, at least on paper, converge with those of the West. In view of the increasingly international scope of Japanese business, the pressures to conform to managerial ideologies widely prevalent in Europe and America will no doubt intensify (Tsurumi, 1976). But whether changes in the structure of Japanese firms or the behavior of their managers to bring them into line with the "efficient" practices of Western corporations are more than surface attempts to acquire legitimacy is uncertain. Some of the recent changes aimed at "rationalizing" Japanese firms do not appear to run very deep. Marsh and Mannari's (1976:140) survey of employees in three Japanese companies, for example, produced a .82 correlation between an employee's job rating and his position in the old seniority system which the newly developed job classification system was designed to replace. Marsh and Mannari speculated that the job classification system was simply the seniority system in disguise.

We find especially intriguing a quite different view of the possibilities for convergence between Japanese and Western organizational forms: that of Dore (1973), which grows out of his conviction that late industrialization explains the exceptional features of the Japanese industrial system. While growing similarity between Japan and the West is what the future holds, Dore contends, it will result from corporations in Europe and the United States increasingly adopting elements of the model on which Japanese enterprise is presently built. These include shifts away from an extreme division of labor in the direction of more flexible occupational roles, greater participation by workers in the decision-making apparatus of the firm, and a broadening of responsibility on the part of the corporation for the well-being and security of its members.

REFERENCES

- Abegglen, James C.
1958 *The Japanese Factory*. New York: Free Press.
- Aldrich, Howard E.
1972 "Technology and organizational structure: a re-examination of the findings of the Aston group." *Administrative Science Quarterly* 17:26-43.
- Azumi, Koya
1978 "Japanese organizations: are they really so different?" Paper presented at the Colloquium Center for Japanese and Korean Studies, University of California, Berkeley.
- Azumi, K. and Charles J. McMillan
1974 "Management strategy and organization structure: a Japanese comparative study." Unpublished paper. Center for Japanese and Korean Studies, University of California, Berkeley.
- Bendix, Reinhard
1956 *Work and Authority in Industry*. New York: Wiley.
- Blau, Peter M. and Richard A. Schoenherr
1971 *The Structure of Organizations*. New York: Basic Books.
- Bohrnstedt, George W.
1970 "Reliability and validity assessment in attitude measurement." Pp. 80-99 in Gene F. Summers (ed.), *Attitude Measurement*. Chicago: Rand McNally.
- Caudill, William and George DeVos
1956 "Achievement, culture, and personality: the case of Japanese-Americans." *American Anthropologist* 56:1102-26.
- Child, John
1972 "Organization structure and strategies of control: a replication of the Aston studies." *Administrative Science Quarterly* 17:163-77.
1973 "Strategies of control and organizational behavior." *Administrative Science Quarterly* 18:1-17.
- Cole, Robert E.
1971 *Japanese Blue Collar: The Changing Tradition*. Los Angeles: University of California Press.
- Connor, John W.
1977 *Tradition and Change in Three Generations of Japanese-Americans*. Chicago: Nelson-Hall.
- Crozier, Michel
1964 *The Bureaucratic Phenomenon*. Chicago: University of Chicago Press.
1970 "The cultural determinants of organizational behavior." Pp. 49-58 in A. R. Negandhi (ed.), *Environmental Settings in Organizational Functioning*. Kent: Comparative Administration Research Institute, Kent State University.
- Donaldson, Lex, John Child, and Howard Aldrich
1975 "The Aston findings on centralization: further discussion." *Administrative Science Quarterly* 20:453-61.
- Dore, Ronald
1973 *British Factory, Japanese Factory: The Origins of National Diversity in Industrial Relations*. Los Angeles: University of California Press.

- Dowdy, Edwin
1973 *Japanese Bureaucracy: Its Development and Modernization*. Melbourne: Cheshire.
- Evan, William M.
1975 "Organizational research across national and cultural boundaries." *International Studies of Management and Organizations* 5:3-113.
- Fuguitt, Glenn V. and Stanley Lieberman
1974 "Correlations of ratios or difference scores having common terms." Pp. 128-44 in H. L. Costner (ed.), *Sociological Methodology, 1973-74*. San Francisco: Jossey-Bass.
- Hickson, David J., C. R. Hinings, C. J. McMillan, and J. P. Schwitter
1974 "The culture free context of organization structure: a tri-national comparison." *Sociology* 8:59-80.
- Hickson, David J., C. J. McMillan, K. Azumi, and D. Horvath
Forth-coming "Grounds for comparative organization theory: quicksand or hard core?" In C. J. Lammers and D. Hickson (eds.), *Organizations Alike and Unlike*. London: Routledge and Kegan Paul.
- Hickson, David J., D. S. Pugh, and Diana Pheysey
1969 "Operations technology and organization structure: an empirical reappraisal." *Administrative Science Quarterly* 14:378-97.
- Inkson, J. H. K., D. S. Pugh, and D. J. Hickson
1970 "Organization context and structure: an abbreviated replication." *Administrative Science Quarterly* 15:318-29.
- Johnson, Richard and William Ouchi
1974 "Made in America under Japanese management." *Harvard Business Review* 5:61-9.
- Kitano, Harry L.
1969 *Japanese Americans: The Evolution of a Subculture*. Englewood Cliffs: Prentice-Hall.
- Kuniyoshi, U.
1977 "Reconsideration of the nature of Japanese lifetime employment." *Organizational Science* 2:12-20.
- McMillan, Charles J., D. J. Hickson, C. R. Hinings, and R. E. Schweck
1973 "The structure of work organizations across societies." *Academy of Management Journal* 16:555-69.
- Maykovich, Minako K.
1972 *Japanese-American Identity Dilemma*. Tokyo: Waseda University Press.
- Marsh, Robert M. and Hiroshi Mannari
1976 *Modernization and the Japanese Factory*. Princeton: Princeton University Press.
- Nakane, Chie
1970 *Japanese Society*. Los Angeles: University of California Press.
- Pascale, Richard Tanner
1978 "Communication and decision making across cultures: Japanese and American comparisons." *Administrative Science Quarterly* 23:91-110.
- Pugh, D. S., D. J. Hickson, C. R. Hinings, and C. Turner
1968 "Dimensions of organization structure." *Administrative Science Quarterly* 13:65-105.
- Pugh, D. S., D. J. Hickson, and C. R. Hinings
1969 "The context of organizational structures." *Administrative Science Quarterly* 14:91-114.
- Rohlen, Thomas P.
1974 *For Harmony and Strength: Japanese White-Collar Organization in Anthropological Perspective*. Los Angeles: University of California Press.
- Schuessler, Karl
1974 "Analysis of ratio variables: opportunities and pitfalls." *American Journal of Sociology* 30:279-97.
- Tracy, Phelps and Koya Azumi
1976 "Determinants of administrative control: a test of a theory with Japanese factories." *American Sociological Review* 41:80-93.
- Tsurumi, Yo-ichi
1976 *The Japanese Are Coming*. Cambridge, Ma.: Ballinger.
- Vogel, E. F. (ed.)
1975 *Modern Japanese Organization and Decision Making*. Los Angeles: University of California Press.
- Yoshino, M. Y.
1968 *Japan's Managerial System: Tradition and Innovation*. Cambridge, Ma.: MIT Press.
1975 "Emerging Japanese multinational enterprises." Pp. 146-66 in E. F. Vogel (ed.), *Modern Japanese Organization and Decision Making*. Los Angeles: University of California Press.
1976 *Japan's Multinational Enterprises*. Cambridge, Ma.: Harvard University Press.

WHITE ENROLLMENT STABILITY AND SCHOOL DESEGREGATION: A TWO-LEVEL ANALYSIS*

MICHEAL W. GILES

Florida Atlantic University

American Sociological Review 1978, Vol. 43 (December):848-864

Recent studies have provided conflicting evidence on the relationship between school desegregation and white enrollment stability. Pettigrew and Green (1976), Farley (1975), Rossell (1976) and Fitzgerald and Morgan (1977) have found desegregation to be unrelated to white enrollment stability. In contrast Bosco and Robin (1974), Lord (1975), Coleman et al. (1975), Munford (1976), Clotfelter (1976) and Giles (1977a; 1977b) have reported declines in white student enrollment concurrent with desegregation. The appropriate query therefore should probably not be whether school desegregation leads to white flight, but instead attention ought to focus on the conditions under which white enrollments decrease as a result of desegregation. Among the most frequently cited correlates of white withdrawal is the level of black concentration. There is disagreement, however, about the structure of this relationship. Some studies have found a linear relationship, whereas others have suggested the presence of a tipping point beyond which white withdrawal accelerates and schools become all-black. The present study reexamines the relationship between percent black enrollment and white enrollment change at both the district and the school levels. The analysis focuses on 60 districts and approximately 1,600 schools located in Southern SMSAs. Higher percent black enrollments are found to be associated with the rate of white withdrawal at both the district and the school levels. In both cases this relationship appears to be curvilinear with white withdrawals increasing exponentially with black enrollments over 30%.

The relationship between school desegregation and white enrollment stability has recently been the subject of considerable scholarly debate. Coleman et al. (1975), in a well publicized study of large school districts, concluded that school desegregation accelerated the exodus of white students from urban schools. The obvious policy implications of this research have stimulated several replications and critiques. Pettigrew and Green (1976) in the most detailed analysis of Coleman's research criticized the sample selection and offered alternative explanations of the results. Furthermore, by augmenting his sample with several large Southern districts they concluded that, contrary to Coleman et al., desegregation

was unrelated to white enrollment declines. Farley (1975), Fitzgerald and Morgan (1977), and Rossell (1976), using slightly different techniques and samples than Pettigrew and Green, arrived at the same conclusion—desegregation was unrelated to white enrollment decline.

While the weight of evidence indicates that declines in white student enrollments are not an inevitable consequence of school desegregation, there is substantial evidence that school desegregation has resulted in white withdrawals in some districts. Munford (1973) and Clotfelter (1976) in separate analyses of Mississippi school districts, report large declines occurring in some districts subsequent to school desegregation. Giles (1977b) reports that in a random sample of 100 Southern school districts, many experienced no change or even gains in white enrollment with desegregation but many others experienced large white enrollment declines. Lord (1975) found that the number of private schools in Charlotte-Mecklenberg, North Carolina, area had almost doubled in the first four years of desegregation, and in some areas of the

* Address all communications to : Micheal W. Giles; Department of Political Science; Florida Atlantic University; Boca Raton, FL 33431.

I would like to thank Krishna Kumar, Douglas S. Gatlin and Lee Sigelman for their helpful comments on an earlier draft of this paper. This article was prepared with support from the National Institute of Education, Contract 400-76-0102. The views expressed are those of the author and should not be ascribed as views of the National Institute of Education.

district as many as 25% of the white students were withdrawn from the public schools. Outside the South, Bosco and Robin (1974) report dramatic declines in white enrollments in the Pontiac, Michigan, public schools after desegregation. Thus, decline in white enrollment appears to be a potential but not an inevitable outcome of school desegregation. The problem confronting social scientists is to specify the conditions which influence white enrollment stability and instability in desegregated schools.

This problem is of more than academic interest. The courts have consistently held that fear of white withdrawal is not sufficient grounds to justify a denial of constitutional right. For example, when the school district asserted in *Monroe v. Board of Commissioners* (391 U.S. 450 at 459) that white students would "flee the system altogether" if desegregation occurred, the Supreme Court reasserted the principle of *Brown* that "... the vitality of these constitutional principles cannot be allowed to yield simply because of disagreement with them." The courts, however, have not been oblivious to the problems of white withdrawals. The Supreme Court declared state legislation unconstitutional which authorized the City of Scotland Neck, North Carolina, to separate itself from the desegregating Halifax County School District (*U.S. v. Scotland Neck*, 407 U.S. 483). This decision was based in part upon the Court's belief that the legislation would create a "refuge for white students of the Halifax County School system." The courts also have shown a willingness to consider the problem of white withdrawals in designing remedies for segregation. For example, a California state court mandated that busing distances in San Diego be less than five miles for fear that longer distances would increase white withdrawals (*Los Angeles Times*, 3/10/77:3).

Sensitivity to the factors influencing white withdrawal and hence the long term stability of desegregation is of even greater importance in light of the U.S. Supreme Court's recent decision in *Pasadena v. Spangler* (427 U.S. 424 [1976]). In *Spangler* the Court held that a Federal District Court could not order a

local school board to readjust its school enrollments on a yearly basis after racially neutral assignment plans had been adopted. In support of its decision the Court cited its previous statement in *Swann v. Charlotte-Mecklenburg* (402 U.S. 1 [1970:32]) that

neither school authorities nor district courts are constitutionally required to make year by year adjustments of racial compositions of student bodies once the affirmative duty to desegregate has been accomplished and racial discrimination through official action is eliminated from the system.

Thus, the District Courts appear to be bound by their initial desegregation decisions and unable to adjust desegregation plans to react to either demographic trends or the reactions of affected parents. In light of the *Spangler* decision, it is of paramount importance that initial desegregation plans be designed to provide long-term stability in racial enrollments. This goal requires that judges and other desegregation planners know the conditions influencing white enrollment stability in desegregated schools.

Among the most frequently discussed correlates of white withdrawal is the level of black concentration. Higher black concentrations are commonly assumed to result in greater white withdrawal. For example, in the Wilmington, Delaware, desegregation case (*Evans v. Buchanan*, 416 F. Supp. 328) the District Court justified the inclusion of the predominantly white Newark School District in the desegregation plan by asserting that "... the stability of any desegregation plan is enhanced by the inclusion . . . of higher white concentrations [355]." Often this relationship is perceived as nonlinear with a tipping or threshold point. Below the threshold, whites are thought to be relatively insensitive to the black concentration. Above the threshold, white withdrawal accelerates with the black ratio until the school becomes all-black. This interpretation is clearly accepted in the Federal District Court's conclusion in *U.S. v. Board of School Commissioners* (332 F. Supp. 655) that "when the percent of black pupils . . . approaches 40 more or less the white exodus becomes accelerated and irreversible."

There is relatively consistent empirical evidence of a linkage between racial concentration and white withdrawal. There is little agreement, however, about the structure of this relationship. Clotfelter (1976), in a study of private school enrollment increases after desegregation in 78 Mississippi counties, found no relationship to the level of black concentration in counties 25% or less black. In counties more than 25% black, private school enrollment increased with the level of black concentration and the increases accelerated in counties more than 55% black. Giles et al. (1975) found that the rate of white withdrawal from desegregated public schools in seven Florida school districts was significantly higher if the school was 30% or more black. The rate of withdrawal, however, did not appear to increase with higher black concentrations above the 30% threshold. In a study of 100 Southern districts, Giles (1977b) found changes in white enrollment to be unrelated to the percent black enrollment in the district until the black ratio exceeded 30%. Above that point, white withdrawal increased with the black concentration. Unlike Clotfelter, Giles found this relationship to be linear above 30%. Munford (1973) in his study of 30 Mississippi districts reports findings consistent with those of Giles. Several studies have reported simple linear relationships between percent black and white withdrawal without a threshold (Farley, 1975; Coleman et al., 1975; Fitzgerald and Morgan, 1977; Lord and Catau, 1977; Stinchcombe et al., 1969). Indeed, in their study of Baltimore, Stinchcombe et al. (1969:134) conclude that "there is no 'tipping point.' Or rather, the 'tipping point' is zero. . . . Once a school is desegregated . . . the proportion Negro is likely to go up each year in a steady fashion. . . ."

In addition to the lack of consistency concerning the structure of the relationship between percent black and white withdrawal, previous studies have suffered several shortcomings. First, such studies have sometimes ignored the source of school desegregation (Coleman et al., 1975; Stinchcombe et al., 1969). School desegregation occurs as a result of governmental intervention and/or as a

product of residential desegregation. When desegregation occurs through government intervention, desegregated schools may be created without disturbing existing patterns of segregated housing. Children are simply transported out of their segregated neighborhoods and placed in desegregated schools. The white parent in this situation is confronted with a single racial stimulus, the percent black in the school. In contrast, when school desegregation results from neighborhood racial succession, the white parent is confronted with at least two racial stimuli, the percent black in the school and the percent black in the neighborhood. These are clearly two different decisional situations for parents and should be different experimental settings for the researcher.

A second limitation of much of the research purporting to deal with white flight is that it focuses solely on the school district as the unit of analysis. Rossell (1976), Farley (1975), Fitzgerald and Morgan (1977), Pettigrew and Green (1976), Coleman et al. (1975) and Munford (1973) all take as their dependent variable decreases in the total white enrollment in school districts. A logical linkage between percent black and white withdrawals, however, is at the level of the individual school. Parents do not reject school districts, they reject schools (Pettigrew and Green, 1976). Furthermore, for designing desegregation plans, knowledge at the individual school level is of particular practical importance. The few studies which have focused on the school level have invariably concentrated on only one district or a very few districts (Wegmann, 1975; Giles et al., 1975; Stinchcombe et al., 1969; Giles, 1977a; Lord and Catau, 1977).

The use of the district as the unit of analysis also has precluded the analysis of some potentially important linkages—most notably the relationship between the percent black in the school, the previous racial status of the school and white withdrawal. A common assumption is that whites will not attend previously black schools. White opposition is thought to arise from the stigma of poor educational quality attached to these schools and their location in predominantly black areas. Indeed, desegregation has often meant the

closing of black schools, although conditions other than fear of white withdrawals have been cited in justification.

It is conceivable that the previous racial status of the school may not only directly affect white withdrawal, but also affect the relationship between the racial concentration in the school and white withdrawal. For example, white withdrawal might be unrelated to percent black enrollment in previously white schools but strongly related to it in previously black schools.

Adopting the individual school as the unit of analysis also allows the effect of variations in school district racial balance on the relationship between school racial balance and withdrawal to be examined. The concentration of blacks in an area consistently has been shown to influence racial climates (Blalock, 1967). We would expect, therefore, that parents with children enrolled in a school 40% black might react quite differently depending upon whether the overall racial balance in the district was 10% black or 60% black. Thus, the findings by Giles et al. (1975) that the rate of rejection of desegregated schools increased dramatically when white children were moved from schools less than 30% black to schools more than 30% black may reflect the racial balances in the districts they studied (18% to 33%). It is conceivable that if the racial balances had been in the 40 to 50% black range the threshold for rejection might have moved upward.

The present study reexamines the linkage between percent black enrollment and white withdrawal at both the school and school district levels. Only districts and schools desegregating under government enforcement are examined. Two questions are addressed at both levels of analysis: (1) Does the rate of white withdrawal vary with the percent black enrollment? (2) If present, is that relationship linear? Three additional questions are examined at the school level: (3) Does the racial status of a school prior to desegregation affect the rate of white withdrawal? (4) Does the percent black in the district influence the rate of white withdrawal from the schools? (5) Does the relationship between the school's racial enroll-

ment and the rate of white withdrawal vary by the racial status of the school before desegregation or by the percent black in the district?

DATA

The analysis focuses on a sample of school districts and their included schools. Eligibility for inclusion in the district sample pool was contingent on four factors. First, only districts included in the Department of Health, Education, and Welfare's racial and ethnic surveys of the public schools in 1967, 1968, 1970, 1972 and 1974 were eligible to be sampled. These surveys provide the student enrollment by race in every school in over 8,000 school districts each surveyed year. Second, the school district must have experienced either court ordered or HEW enforced school desegregation within the 1968 through 1970 time period. Limiting the analysis to districts experiencing government induced desegregation minimizes the confounding effects of residential desegregation.¹ Furthermore, focus on those districts desegregating in 1968, 1969 or 1970 allows for information on the racial balances in the schools for at least one point in time prior to desegregation, 1967, and for at least two points in the postdesegregation period, 1972 and 1974. Third, since virtually all government induced desegregation between 1968 and 1970 occurred in the South, the study was confined to that region. Fourth, only school districts located in metropolitan areas were eligible for inclusion in the sample. Residential choice and mobility are greater in urban than in rural settings. We would expect, therefore, that white enrollment stability would be more of a problem in the former than in the latter.

The sample was drawn by first constructing a list of all school districts located in Southern SMSAs as defined in the 1970 census. Second, all districts for which data could not be found in the 1967 through 1974 HEW surveys were excluded from the list. Third, the Office of

¹ The effects of residential desegregation are not eliminated by focusing on districts experiencing government induced desegregation.

Civil Rights (OCR) of the Department of Health, Education and Welfare assembled a list of the timing and mode of desegregation for over 1,400 Southern school districts. Using this information, we deleted districts not reported as desegregating under court order or HEW plans in either 1968, 1969 or 1970 from the sample pool. One hundred school districts were then sampled from the pool. The 50 largest school districts were included automatically in the sample. An additional 50 were drawn by stratified random sampling to obtain variation in the district racial balance.

Accurate determination of the timing of desegregation in each district is essential to the study design. If the OCR date for desegregation in a district is incorrect, changes in racial balances arising from implementation of the desegregation plan will be confused with other sources of enrollment instability. To assure the accuracy of the date of desegregation, we contacted each of the 100 selected school districts by telephone and questioned them concerning the timing of desegregation. This information was then compared with the OCR date for desegregation and with the pattern of enrollment stability by race in the district. Desegregation should be accompanied by alteration in the racial balances in the schools of a district. In particular, we would expect to see evidence of blacks attending previously all-white schools and whites attending previously all-black schools. For 60 of the sampled school districts, information from all three sources indicated the same date for desegregation. For the remaining 40 there was disagreement among the information sources. For 17 of these 40, two of the three sources agreed on the desegregation date. In most cases this involved the enrollment data supporting either the local school authorities' or the Office of Civil Rights's estimate of the date of desegregation. When agreement between two sources could be obtained, the date was accepted as the correct one for desegregation. In most cases, this involved a change of one year (e.g., the OCR said a district desegregated in 1968 but the district and the school enrollment data indicated that the change occurred in 1969). Conversa-

tions with the school officials suggest that these differences may have arisen from the OCR reporting the first year of a two-year desegregation plan as the year of desegregation or the year of its adoption rather than the year of its implementation. As a result of these changes 15 districts were found to be outside the 1968-70 time frame for the study. These districts along with the 18 districts for which all three sources differed over the timing of desegregation are excluded from the study. The resulting sample consisted of six districts that desegregated in 1968, 24 that desegregated in 1969 and 36 that desegregated in 1970. For simplicity, only the analysis for districts and schools desegregating in 1969 or 1970 are presented below.²

The 60 districts included in the analysis averaged about 24,000 students. The smallest district enrolled slightly less than 1,000 students and the largest district enrolled about 240,000. The districts averaged approximately 30% black in enrollment. This relatively high average reflects the sampling stratification which assigned a higher probability of inclusion to school districts with large black concentrations. Fourteen of these districts desegregated under HEW enforcement and the remainder under court order.

Data on enrollment by race in each school in the sampled districts were drawn from the HEW surveys for 1967, 1968, 1970, 1972 and 1974. Approximately 2,000 schools were included in these districts. Those schools for which enrollment data could not be identified for every surveyed year were omitted from the analysis. This resulted in a sample of approximately 1,600 schools. The loss of 400 schools is consistent with HEW's estimate of 10 to 15% error in the assignment of school identification codes on a year to year basis.

The dependent variable in the study is the rate of white enrollment change. This is operationalized as the percent change in white enrollments between two surveyed years:

$$\frac{W_t - W_{t-2}}{W_{t-2}}$$

² For a list of districts see the Appendix.

where W_t = the white enrollment in a surveyed year, W_{t-2} = the white enrollment two years before W_t . If white enrollments decrease this index will assume a negative value and if white enrollments increase the index will become positive. If no change in white enrollments occurs, the index will be zero.

Four measures on the dependent variable are available for each school district—the percentage change in white enrollment between 1967 and 1968 ($[DWE_{1968} - DWE_{1967}]/DWE_{1967}$), between 1968 and 1970 ($[DWE_{1970} - DWE_{1968}]/DWE_{1968}$), between 1970 and 1972 ($[DWE_{1972} - DWE_{1970}]/DWE_{1970}$), and between 1972 and 1974 ($[DWE_{1974} - DWE_{1972}]/DWE_{1972}$). The first measure provides an estimate of white enrollment stability just prior to desegregation. The second measure provides a comparison of white enrollments before and after desegregation. The third and fourth measure white enrollment stability for time periods after desegregation.

At the school level only two measures on the dependent variable are available—the percentage change in white enrollment between 1970 and 1972 ($[SWE_{1972} - SWE_{1970}]/SWE_{1970}$) and between 1972 and 1974 ($[SWE_{1974} - SWE_{1972}]/SWE_{1972}$). White withdrawals from schools also may occur with the onset of desegregation and result in higher black enrollments than expected by the desegregation plan. An examination of white refusals to attend desegregated schools in the first instance necessitates a comparison between the expected black/white enrollments and the actual black/white enrollments in each desegregated school (Giles, 1977a). For example, if a district desegregated in the fall of 1970, the white enrollment projected for each school under the desegregation plan would be compared to the number of whites actually enrolled in that school in 1970. Reliance on HEW school census data limits the present study to consideration of white enrollment changes after the implementation of desegregation has occurred in a district. Thus, the phenomena of white withdrawal with the onset of desegregation is not considered at the school level in this study. However, it would seem reasonable to expect that the

factors which influence white enrollment stability in desegregated schools also would influence, in similar ways, white decisions to attend or not to attend newly desegregated schools.

In schools with low white enrollment even small changes in the number of white students results in high percentage changes. For example, if a school has ten white students enrolled at t_1 and five white students enrolled at t_2 , the percent white enrollment change is $-.50$ ($[5-10]/10$). In contrast, a school with 100 whites enrolled at t_1 must lose 50 whites by t_2 to have a comparable percent white enrollment change. To compensate for this problem, we excluded all schools enrolling fewer than 50 white students from the school level analysis.³ This does not eliminate the problem but substantially reduces its magnitude.

The black concentration in the schools and districts is measured straightforwardly as the percent black enrollment in each. The previous status of the schools is measured by their percent black enrollment in 1967. Some of the schools in the sample districts desegregated prior to 1968 probably as a result of modest redistricting or "freedom of choice" desegregation plans. However, the overwhelming majority of the schools were either predominantly white (less than 20% black) or predominantly black (more than 90% black). For analysis purposes, previous racial status is coded as two dummy variables. The first is coded one if a school was integrated in 1967 and zero otherwise. The second is coded one if a school was predominantly black in 1967 and zero otherwise.⁴

ANALYSIS

The relationships between percent

³ Preliminary analysis also was conducted for schools with white enrollment greater than or equal to 100. The results were comparable to the analysis for schools with white enrollments greater than or equal to 50. A few schools with erratic enrollment patterns also were deleted.

⁴ With both variables included in a regression equation the first contrasts percent white enrollment change in previously black and previously white schools and the second contrasts percent white enrollment change in integrated and previously white schools.*

white enrollment change and the independent variables are examined by means of regression analysis. In this analysis the percent white enrollment change between two surveyed years is regressed on the value of the independent variables in the first year of the comparison. For example, the percent black school enrollment in 1970 is used to predict percent white enrollment change between 1970 and 1972, and percent black school enrollment in 1972 is used to predict percent white enrollment change between 1972 and 1974. Analyses are conducted separately at the district and school levels.

District-Level Analysis

Previous studies have reached varying results with regard to the linkage between desegregation and white withdrawal. The white enrollment patterns in the districts studied appear to support such a linkage. On average white enrollments in these districts were stable prior to desegregation but decreased subsequent to the implementation of desegregation (Table 1). However, only two of these declines are statistically significant, and all are substantively trivial in magnitude. Furthermore, the districts varied considerably around these means. In the 1968-70 time period, for example, one district increased its white enrollment by 15% while another declined by 37%. This variability supports the conclusion that white withdrawals are not an inevitable result of school desegregation.

The extent to which variation in white enrollment changes among the districts is related to their percent black enrollment is examined in Table 2.⁵ The 1967-68 period

Table 1. District Level Mean Percent White Enrollment Change by Time Period

Time Period	Mean
1967-68	.003
1968-70	-.030*
1970-72	-.039**
1972-74	-.013

* Significant at .05.

** Significant at .01.

precedes desegregation in these districts. In Table 1 we saw that on average little change in white enrollment occurred during this time period. In Table 2 it appears that those changes that did occur were unrelated to the black concentration in these districts (eq. 2.1). The percentage change in white enrollment between 1968 and 1970 provides a before and after desegregation comparison. The black concentration in the districts is strongly related to the percent white enrollment change in this time period.⁶ Approximately half the variation in percent white enrollment change is explained by the percent black enrollment. As expected, the effect of percent black is negative with higher black concentrations resulting in declines in white enrollment. The pattern set in the 1968-70 time period continues in the 1970-72 time period. The relationship between percent black enrollment and white withdrawal appears somewhat diminished in the 1970-72 period but remains statistically significant. In the 1972-74 time period white enrollment declines continue to be significantly associated with percent black enrollment, but the regression coefficient and the variance

⁵ The regression analysis was conducted with three outlying districts removed: Conroe, Montgomery; Jasper, Walker; and Crawfordsville, Crittenden. The first two districts had relatively low black enrollments (8 to 15%) and experienced large increases in black enrollment between 1968 and 1970. Indeed, white enrollment in the Jasper school district increased by 40% between 1968 and 1970. In contrast, Crawfordsville school district was approximately 80% black and lost about 66% of its white enrollment between 1968 and 1970. These gains and losses are consistent with a relationship between percent black and white enrollment decline. However, the values for these districts departed to such a degree from the

other districts that a more accurate picture of that relationship was provided by deleting them from the analysis.

⁶ Since both the dependent variable, percent white enrollment change, and the independent variable, percent black enrollment, contain the term total white enrollment, the presence of a relationship between the two variables might be a mathematical result of this common term. To examine this possibility the regression results at both the district and the school levels were controlled for total white enrollment. The results of the analysis were not altered by these controls. Thus, the relationships found between percent black and percent white enrollment change do not appear to be artifacts of the way in which the variables are defined.

Table 2. Unstandardized Regression Coefficients for Percent Black and Percent Black Squared on District Percent White Enrollment Change by Time Period

Time Period	Intercept	% Black	(% Black)*	R ²
1967-68				
eq. 2.1	.0175	-.0462		.00
2.2	-.0020	+.0974	-.2005	.01
1968-70				
eq. 2.3	.0962	-.3977**		.48
2.4	.0348	+.0535	-.6241*	.52
1970-72				
eq. 2.5	.0741	-.3464**		.37
2.6	.0348	+.3565	-.9258**	.49
1972-74				
eq. 2.7	.0318	-.1341*		.08
2.8	.0298	-.1211	-.0158	.08

* Significant at .05.

** Significant at .01.

explained by percent black enrollment are much smaller than in either the 1968-70 or 1970-72 time periods.

Clotfelter (1976), Munford (1973) and Giles (1977b) found nonlinear relationships between percent black and white withdrawal. Specifically, they found that whites were insensitive to racial concentrations of less than 25 (Clotfelter, 1976) to 35% (Giles, 1977b) black. A common procedure for testing the linearity of a relationship between a dependent and an independent variable in a regression equation is to square the independent variable, add it to the equation and test its significance. The results of this analysis for the relationship between percent black district enrollment and percent white enrollment change are presented in equations 2.2, 2.4, 2.6 and 2.8. In both the 1968-70 and 1970-72 time periods percent black enrollment squared is significant, indicating the presence of a nonlinear relationship.⁷ The regression lines fitted by

these equations are presented in Figure 1 for the 1968-70 period and Figure 2 for 1970-72. The relationships revealed in Figures 1 and 2 closely resemble the findings of Clotfelter (1976), Munford (1973) and Giles (1977b). Among districts less than approximately 30% black, differences in the percent black enrollment are not consistently related to differences in the percent white enrollment change in either time period. Indeed, for over 90% of these districts the change in white enrollment fell within a range of $\pm 10\%$ from zero. In contrast, among districts above 30% black, the rate of white withdrawal increases exponentially with higher percent black enrollments.

While the equations for the two time periods differ somewhat, the estimated percent white enrollment changes in the two time periods are virtually identical. For example, a school district with a 75% black enrollment in 1968 is predicted to lose approximately 28% of its white students between 1968 and 1970. Similarly, a district that is 75% black in 1970 is predicted to lose approximately 28% of its white enrollments between 1970 and 1972. Of course, a district that is 75% black in 1968 would be approximately 80% black in 1970, having lost 28% of its white enrollment and assuming no increase in its black enrollment. Given an 80% black enrollment, a 33% decline in white enrollments between 1970 and 1972 is projected. In total such a district is predicted to lose more than half of its white enrollment between 1968 and 1972.

The pattern of accelerating losses between 1968 and 1972 conforms to the expectations of the tipping model. However, the tipping model would predict this pattern to continue until the districts became all-black. This does not occur. In the 1972-74 time period percent black enrollment squared is not significant (eq. 2.8), indicating that the relationship between percent black and the rate of white withdrawal is linear in this time period (eq. 2.7). Consistent with the findings for 1968-70 and 1970-72, equation 2.7 pre-

⁷ If the deleted districts are included (see fn. 5), the quadratic term is nonsignificant for the percent change in white enrollment between 1970 and 1972. This is largely due to the effect of the Crawfordsville school district. As a result of a decline of approximately 66% in white enrollment between 1968 and 1970, only 61 whites were enrolled in the districts in 1970. An additional 26 white students were added

between 1970 and 1972. This constitutes a 43% increase in white enrollment for a district over 90% black in 1970.

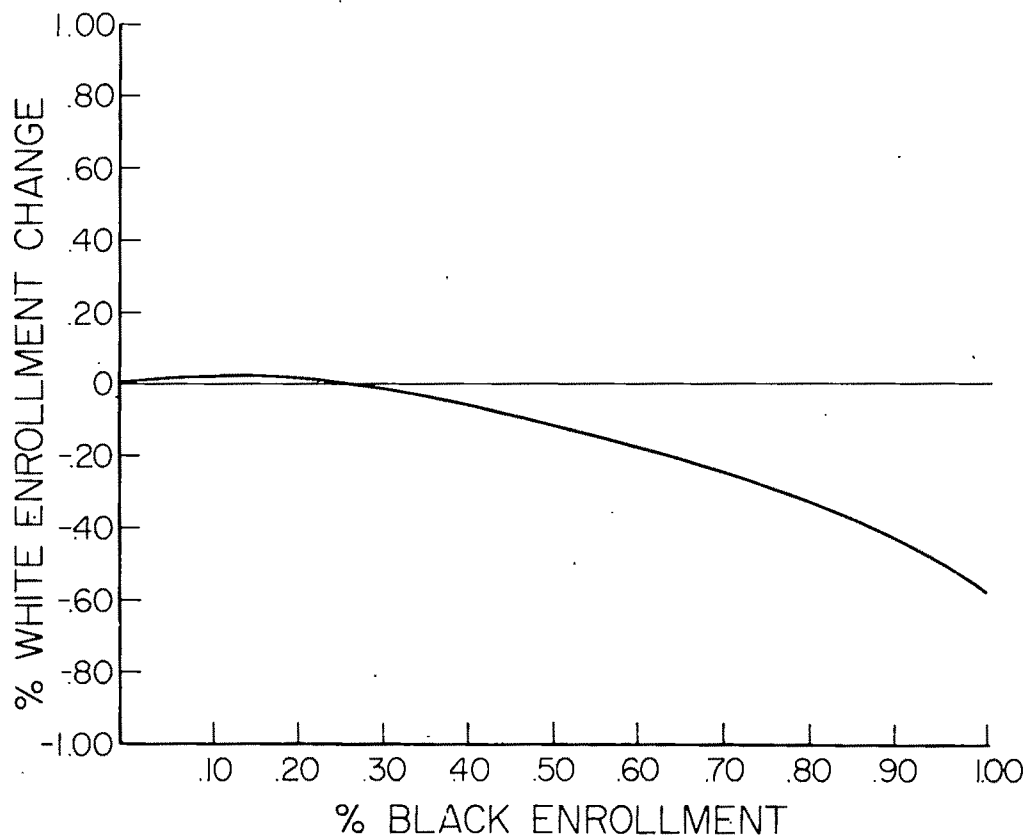


Figure 1. Estimated District Percent White Enrollment Change between 1968 and 1970

dicts that higher black enrollments will yield higher rates of white withdrawal only among districts more than approximately 25 to 30% black.⁸ Among these districts, however, the estimated percent white enrollment decline is far smaller than in the previous time periods. For example, a district that is 75% black in 1972 is predicted to lose only about 7% of its white enrollment between 1972 and 1974. Thus, rather than accelerated, the rate of white withdrawal is actually diminished between 1972 and 1974. Apparently, most parents who were sensitive to the black concentration in their district and who could afford to, withdrew their children either with the onset of desegregation or in the first year or two after its

implementation. While some districts continue to experience declines in white enrollments during the 1972-74 time period, these declines result largely from factors other than the level of black concentration.

Pettigrew and Green (1976) found that losses in black enrollment were also associated with higher percent black enrollments among their districts. They suggest that percent black is a surrogate variable for several unfavorable characteristics of a district which result in both white and black enrollment declines. Such an interpretation is not supported by the present data. Percent black enrollment is not significantly related to changes in black enrollment among the study districts.

In summary among the sampled districts percent white enrollment decline is associated with percent black enrollment in three of the four time periods examined. The exception is the 1967-68 time period

⁸ Equation 2.7 predicts that a district 24% black will lose .04% ($+.0218 - [.1341 \cdot .24]$) white enrollment between 1972 and 1974. Below 24%, equation 2.7 predicts smaller percent white gains with increases in percent black enrollment.

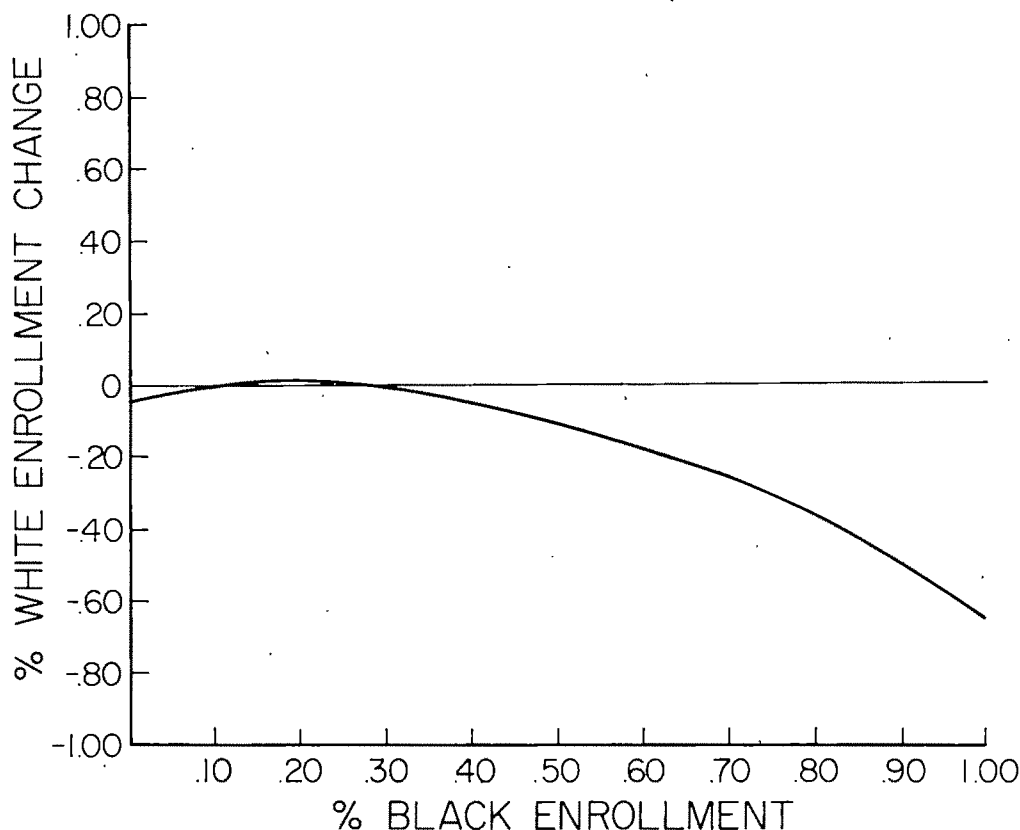


Figure 2. Estimated District Percent White Enrollment Change between 1970 and 1972

which precedes widescale desegregation in these districts. In the 1968-70 and 1970-72 time periods percent change in white enrollment was relatively minor and unassociated with the level of black concentration among districts with enrollments less than 30% black. Among districts more than 30% black, the rate of white withdrawal increased exponentially with increases in the percent black enrollment. Higher black concentrations also resulted in greater percent white withdrawal between 1972 and 1974 in districts more than 30% black, but the relationship was linear not exponential, and the predicted declines were far less than in the 1968-72 time period. This result suggests that the effect of percent black on white withdrawal diminishes over time, contrary to the expectations of the tipping model.

The value of these findings for policy purposes depends on their generalizability beyond the districts included in the study

sample. To assess the generalizability of the findings at the district level, we used equation 2.2 to estimate the percent white enrollment changes between 1968 and 1970 in an independently drawn random sample of 100 Southern school districts which desegregated in 1969 or 1970.⁹ This sample differed from the study sample in several ways. The districts were smaller, averaging approximately 6,500 in enrollment, more densely black, averaging approximately 40% black, and were rural as well as urban in location. Despite these differences, for approximately 65% of the districts the actual percentage change in white enrollment fell within $\pm .10$ of the estimates. This constitutes a 35% improvement in prediction over estimates based simply on the mean percent white enrollment change in our study districts.¹⁰

⁹ These districts are the same as those analyzed by Giles (1977b).

¹⁰ Estimating that white enrollment in the supplemental district would equal the mean white enroll-

Table 3. Unstandardized Regression Coefficients for Percent White Enrollment on School Percent Black Enrollment, District Percent Black Enrollment and Previous Status of the School for Schools in Districts Desegregating in 1969 or 1970 (N=1,368)

	Inter- cept	School % Black	(School % Black)*	District % Black	Previous Status		R ²
					Black	Integrated	
1970-72							
Equation 3.1	-.085	-.067*					.01
Equation 3.2	-.112	+.319**	-.555**				.02
Equation 3.3	-.025			-.266**			.03
Equation 3.4	-.097				.040	-.142**	.02
Equation 3.5	-.053	.417**	-.657**	-.268**	.093**	-.111**	.07
1972-74							
Equation 3.6	-.044	-.202**					.03
Equation 3.7	-.088	.345**	-.737**				.08
Equation 3.8	.008			-.341**			.04
Equation 3.9	-.079				-.070	-.103	.01
Equation 3.10	-.031	.410**	-.739**	-.253**	.016	-.028	.10

* Significant at .05.

** Significant at .01.

Thus, the results of the present study provide at least rough guidance about what to expect in terms of white withdrawals at the district level when desegregation occurs.

School-Level Analysis

The relationships between changes in school enrollment and the independent variables are shown in Table 3. In both the 1970-72 and 1972-74 time periods, percent black enrollment is significantly and negatively linked to white enrollment stability (eqs. 3.1 and 3.6). Schools with high black enrollments at the beginning of each comparison period experience larger declines in white enrollments over the two year periods than do schools with low black enrollments. To test the linearity of these relationships, we squared the percent black enrollment in each school and added it to the regression equations. The results of this procedure are presented in equations 3.2 and 3.7. For both time periods percent black enrollment squared is statistically significant, indicating that the

relationship between percent black school enrollment and white enrollment change is nonlinear. The shape of these relationships is presented in Figures 3 and 4.

For the 1970-72 time period equation 3.2 predicts that up to about 30% black, higher black concentrations actually are associated with smaller losses in white enrollment. Above 30% black, the rate of white withdrawal increases exponentially with higher black enrollments. A similar pattern appears for the 1972-74 time period although the downward slope appears to commence between 20 and 30% black. The relationship between percent black and percent white enrollment change also appears to be slightly stronger in the 1972-74 period. A school that is 75% black in 1970 is projected to lose approximately 18.5% of its white enrollments between 1970 and 1972, whereas a school that is 75% black in 1972 is estimated to lose 24% of its white enrollments between 1972 and 1974. Over the four year period a school that is 75% black in 1970 is projected on average to lose approximately 40% of its white students. While the relationship is significant in both time periods, only a small percentage of the variance in white enrollment change is explained by the percent black enrollment. Thus, the fit of the data to the projections portrayed in

ment in the study districts $\pm .10$ yields 54 errors. Estimates based on equation 2.2 $\pm .10$ produced 35 errors. This amounts to a 35% reduction in errors ($(54-35)/54 = .35$).

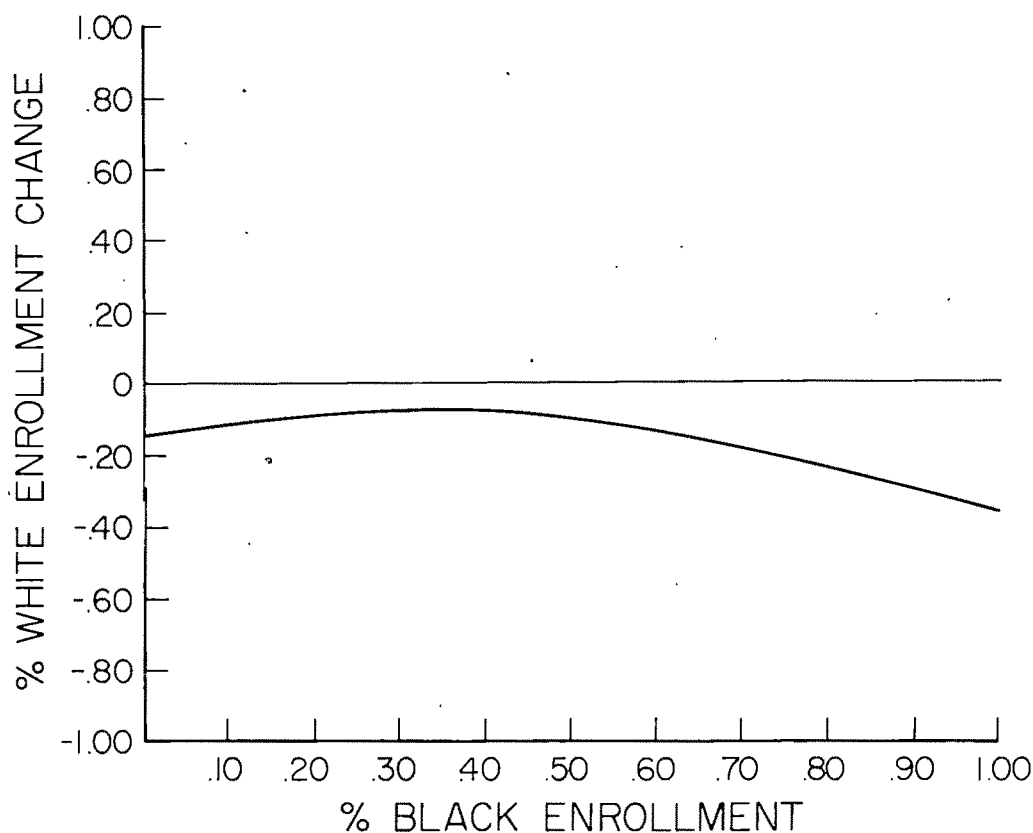


Figure 3. Estimated School Percent White Enrollment Change between 1970 and 1972

Figures 3 and 4 is imprecise and, indeed, some high percent black enrollment schools may experience little decline in white enrollment.

The relationship between percent black enrollment and percent white enrollment change at the school level differs only slightly in the 1970-72 and 1972-74 time periods. In contrast, at the district level the relationship between percent black enrollment and percent white enrollment change was found to be diminished in the 1972-74 period. This difference in results suggests that white withdrawals may continue in response to higher percent black school enrollments even as the effect of percent black enrollment on white withdrawals at the district level subsides.

The relationship between the percent black enrollment in the district and percent change in white school enrollment is shown in equation 3.3 for the 1970-72 period and equation 3.8 for the 1972-74 period. In both time periods percent black in

the district is significantly and negatively related to white enrollment change. The magnitude of the effect however is not great. In both time periods a 10% higher black enrollment in the district yields, on average, about a 2.5% greater decline in white enrollment once the effects of the other independent variables are taken into account (eqs. 3.5 and 3.10). As might be expected, percent black school enrollment and percent black district enrollment are positively correlated ($r = .38$ for 1970-72 and $r = .40$ for 1972-74). The apparent effect of higher percent black school enrollments on white enrollment stability might simply reflect the greater likelihood of such schools being located in higher percent black districts (or vice versa). However, this is not the case. In equations 3.5 and 3.10 the effect of each independent variable is shown net of the effect of all other independent variables. Both the percent black enrollment in the school and in the district remain significant de-

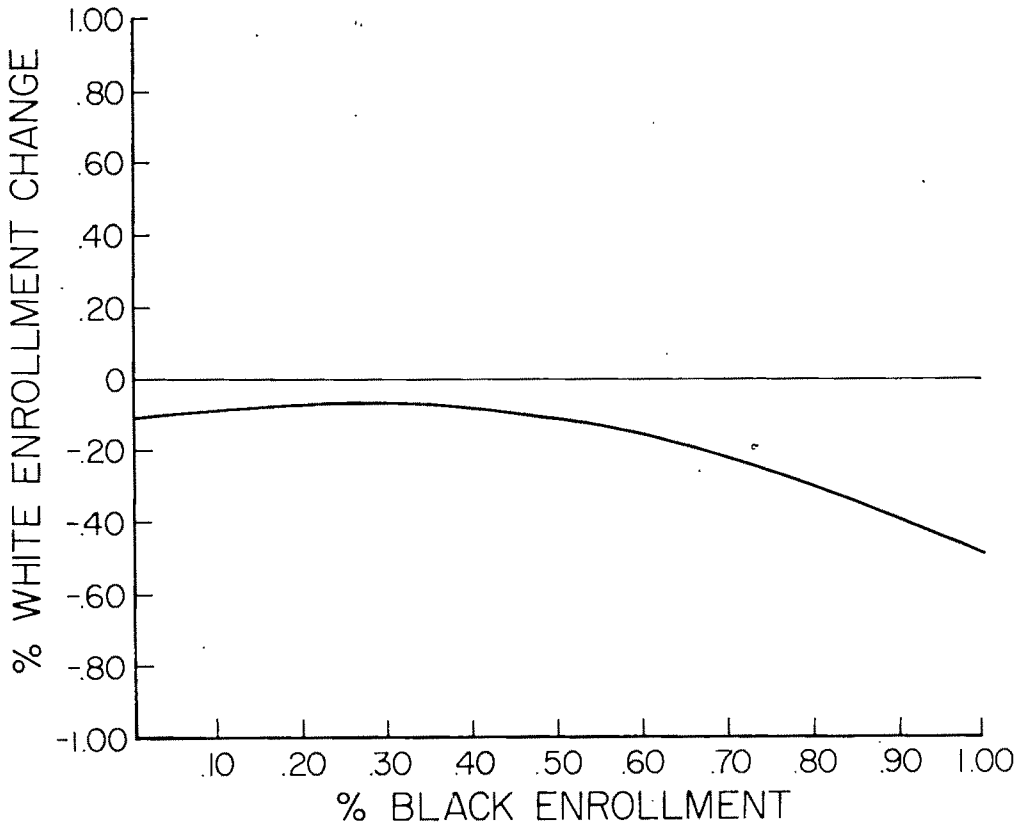


Figure 4. Estimated School Percent White Enrollment Change between 1972 and 1974

spite these controls. Pettigrew and Green (1976) assert that parents reject schools not school districts. These results indicate that, to the contrary, the rate of white withdrawal is influenced by the level of black concentration in both the school and the school districts. Regardless of their percent black enrollments, schools in high percent black districts lose more white student than do schools in low percent black districts and, regardless of the percent black enrollment in the district, high percent black schools lose more white students than do low percent black schools.

The effect of the previous racial status of the school on white enrollment stability is shown in equations 3.4 and 3.9. The pattern is erratic. In the 1970-72 period, schools that were integrated prior to 1968 experienced significantly greater white withdrawals than did schools that were previously white. On average, the losses were 14% greater before (eq. 3.4) and 11%

greater after (eq. 3.5) controlling for the effects of the school and district percent black enrollments. In contrast, previously black schools actually gained white students between 1970 and 1972. It must be assumed that efforts to desegregate black schools lingered on in the 1970-72 time period and that the increases in white enrollments in these schools result from the reassignment of white students. In the 1972-74 time period the effects of previous status disappear. Neither previously integrated nor previously black schools differ significantly from the previously white schools in percent white enrollment changes in this period, once the effects of the percent black enrollments in the school and the district are controlled (eq. 3.10).¹¹

Does the relationship between the per-

¹¹ If the previous status variables are deleted from equations 3.5 and 3.10, the regression coefficients for the remaining variables are virtually unchanged.

cent black enrollment in the school and white enrollment change vary by the black concentration in the district and/or the previous racial status of the school? To address this question, we constructed interaction terms by multiplying the squared percent black school enrollment by each of the remaining independent variables. The resulting three terms were added to equations 3.5 and 3.10. For the 1970-72 time period none of these terms was statistically significant. This result suggests that the effect of the percent black enrollment in a school on white withdrawal was the same during this period regardless of the previous status of the school or whether the district was high or low in black concentration. In the 1972-74 period, one interaction term proved to be significant, percent black in the school by percent black in district. The sign of the coefficient was positive indicating that white withdrawal from high percent black schools located in high percent black districts would be less than predicted by the simple additive model (eq. 3.10). Close analysis suggests that this effect arises from a nonlinear relationship between percent white enrollment change and district percent black enrollment.¹² Approximately a dozen schools located in districts more than 70% black in 1972 lost sizeable numbers of whites in the 1970-72 period but then became relatively stable in the 1972-74 period. With these schools removed, the interaction effect between percent black enrollment in the school and percent black enrollment in the district is no longer statistically significant.

In summary, among the schools examined the percent black enrollment is negatively related to percent white enrollment between 1970 and 1974. This relationship is nonlinear. Among schools less than 20 to 30% black, increasing black enrollment bears little relationship to percent white enrollment change. Among schools more than 20 to 30% black, increases in percent

black enrollment yield increasingly larger percentage declines in white enrollment. With the exception of schools in a few districts with high percent black enrollment which appear to have stabilized their white enrollment in the 1972-74 period, the relationship between percent black school enrollment and percent white school enrollment change does not vary by the percent black in the district or the previous status of the school.

The relationship between the previous racial status of the school and white withdrawal is erratic. Schools integrated prior to desegregation appear to experience white withdrawal in the 1970-72 time period while previously black schools experienced gains. The latter appear to result from continued desegregation efforts. By the 1972-74 period, previous status is unrelated to enrollment change.

Percent black at the district level also is negatively related to changes in white enrollment at the school level. Schools located in high percent black districts experience greater percentage declines in white enrollment than do schools in low percent black districts. As previously noted, a departure from this pattern appears among high percent black districts in the 1972-74 time period. The relationship between percent black district enrollment and percent white enrollment change at the school level differs somewhat from the findings at the district level. At the district level, the relationship between percent black enrollment and percent white enrollment change was nonlinear in the 1970-72 period, but at the school level the effect appears to be linear for the same time period. Results at the district and school levels may differ for at least two reasons. First, the school-level analysis is sensitive to enrollment shifts among the schools in a district as well as to movement to private schools and to other school districts. The district-level analysis only focuses on enrollment shifts to private schools and other school districts. Second, not all the schools in the sampled districts are included in the analysis. Schools for which data could not be found for each surveyed year between 1967 and 1974 were deleted from the study. These two factors may combine to

¹² Tests of the linearity of the relationship between all the independent variables and percent white enrollment decline were conducted as a preliminary step in the analysis. The nonlinearity of the relationship between school percent white enrollment change and district percent black was discovered in this preliminary analysis.

produce slightly different results at the school and district levels. For example, high percent black districts may have closed schools as a result of declining white enrollments. The losses experienced by the closed schools would not appear in the school-level analysis. Furthermore, declines in white enrollments in other schools would be offset by the additional whites transferred from the closed schools. Under such circumstances the school-level analysis would show smaller white enrollment declines than would the district-level analysis.

The results of the school-level and district-level analyses also differ in the amount of variance explained. None of the variables taken either individually or in combination explains more than 10% of the variance in percent white enrollment change at the school level. This difference is not totally unexpected. In general disaggregated data will yield lower explained variance even when relationships are the same. Nevertheless, the low explained variance indicates that changes in white enrollments at the school level are largely a result of factors other than those included in the present study.

The policy significance of this low explained variance is dramatically illustrated when an attempt is made to generalize the findings beyond the present sample. Using equation 3.5, we computed estimates of the percent white enrollment change between 1972 and 1974 for approximately 600 schools in 11 districts desegregating in 1971 and 1972.¹³ For approximately 50% of the schools the actual percent white enrollment change fell within $\pm .10$ of the estimates. However, approximately the same level of accuracy was achieved by simple predicting for each school the mean percent white enrollment change as occurred in the study schools. The projections in Figures 3 and 4, thus, provide little guidance for estimating changes in school enrollments. On the other hand, a separate analysis of the relationship between percent white enrollment change

and percent black among these 600 schools found that relationship to be curvilinear with the rate of white withdrawal increasing exponentially above 30% black. The consistency of these results cannot be ignored.

DISCUSSION

Recent studies have produced conflicting evidence on the relationship between school desegregation and white enrollment declines. Some studies have found desegregation to be unrelated to changes in white enrollment, while others have clearly linked school desegregation to white enrollment declines. Therefore the appropriate query is probably not whether school desegregation leads to white withdrawal but rather, what factors influence white enrollment stability and instability in desegregated schools.

The results of the present analysis indicate that, at least at the district level, percent black enrollment is one such factor. Districts less than roughly 30% black experienced only moderate white enrollment instability which was unrelated to the level of black concentration. Among districts above 30% in black enrollment, the rate of white withdrawal increased exponentially with increases in percent black enrollment. These results suggest that, on average, school districts with less than roughly 30% black enrollment can be desegregated without experiencing drastic declines in white enrollment. For districts with enrollments more than 30% black, the policy implications of the study vary with the black enrollment. Among districts only slightly over 30%, the rate of white withdrawal is predicted to be small and should constitute little hinderance to desegregation. On the other hand, majority black districts are predicted to experience significant rates of white withdrawal. In such districts the results support the need for multiple district desegregation plans. By combining high percent black districts with surrounding low percent black districts, the overall black concentration can be reduced and the problem of white withdrawal reduced correspondingly. Failing the adoption of multiple district plans or other successful com-

¹³ These 11 districts were inadvertently included in the original sample and data on their school enrollments were collected prior to discovery of the fact that they fell outside the study time-frame.

pensatory strategies, desegregation of districts with high percent black enrollments is predicted to be costly in terms of white withdrawals.

The relationship between percent black enrollment and white withdrawals at the school level is less clear. Percent black enrollment was found to be significantly related to white enrollment stability. The relationship was curvilinear and resembled closely in form the results at the district level. Despite the consistency of these findings, only a small percentage of the variance in white enrollment at the school level was explained by the percent black enrollment. Stated in practical terms, the actual percent white enrollment changes varied considerably around the estimates based on the percent black enrollment. Does this mean that percent black enrollment can be ignored in desegregation planning? We think not. Given the consistency of the results, a prudent policymaker would be wise to avoid majority black schools when possible.

Three limitations of the present study should be kept in mind by researchers and policymakers alike. First, at the school level the study focuses on stability after desegregation has occurred. If parents refuse to allow their children to attend previously black schools or schools planned to have high racial balances in the initial year of desegregation, then the effect of these variables is overlooked in the present study. Giles (1977b) finds just such a pattern for previously black schools in one Southern metropolitan school district. Previously black schools experienced significantly lower white enrollments than planned in the first year of desegregation but the enrollments were stable at that lower level in the second and third years of implementation. Given this limitation, policymakers would be premature to assume from these findings that parents are insensitive to the racial status of a school or its black concentrations.

Second, the districts examined in the present analysis experienced government induced desegregation. The predictions, therefore would be modified for districts where residential desegregation is occurring or where whites are leaving for other

reasons (e.g., general trends toward suburbanization). Third, the research has focused on Southern districts and schools and in the strictest statistical terms is only generalizable to that region. Until sufficient desegregation occurs outside the South to provide a basis for research, however, reliance on the experience of the South would appear the only alternative to speculation.

APPENDIX

Districts Included in the Study by State and County

<i>State</i>	<i>District</i>	<i>County</i>
Alabama	Athens	Limestone
	Attalla	Etonah
	Baldwin	Baldwin
	Bessemer	Jefferson
	Birmingham	Jefferson
	Gadsden	Etowah
	Jasper	Walker
	Limestone	Limestone
	Montgomery	Montgomery
	Phenix	Russell
	Russell	Russell
	Walker	Walker
Arkansas	Altheiner	Jefferson
	Crawfordsville	Crittenden
	Dollarway	Jefferson
	Earle	Crittenden
	Texarkana	Miller
Florida	Turrell	Crittenden
	Alachua	Alachua
	Dade	Dade
	Escambia	Escambia
	Leon	Leon
Georgia	Orange	Orange
	Dougherty	Dougherty
	Fulton	Fulton
	Walker	Walker
Louisiana	Bossier	Bossier
	Caddo	Caddo
	Calcasieu	Calcasieu
	Lafayette	Lafayette
	St. Bernard	St. Bernard
Mississippi	St. Tammany	St. Tammany
	Rankin	Rankin
North Carolina	Ashboro City	Randolph
	Cumberland	Cumberland
	Durham City	Durham
	Durham	Durham
	Guilford	Guilford
	Orange	Orange
	Wake	Wake

- | | | | |
|----------------|--|--|--|
| South Carolina | Berkeley
Charleston
Pickens | Berkeley
Charleston
Pickens | Fitzgerald, Michael and David R. Morgan
1977 "Assessing the consequences of public policy: school desegregation and white flight in urban America." Paper presented to the annual meeting of the Mid-West Political Science Association, Chicago. |
| Texas | Chapel Hill
Cleveland
Conroe
Houston
Lamar
Liberty
Lubbock
Mart
Port Arthor
Richardson
San Angelo
South Park
Wilmer Hutchins | Smith
Liberty
Montgomery
Harris
Fort Bend
Liberty
Lubbock
McLennan
Jefferson
Dallas
Tom Green
Jefferson
Dallas | Giles, Micheal W.
1977a "Racial stability and urban school desegregation." <i>Urban Affairs Quarterly</i> 12:499-510.
1977b "School desegregation and white withdrawal: a test of the tipping-point model." Mimeo. Department of Political Science, Florida Atlantic University.
Giles, Micheal W., Everett F. Cataldo and Douglas S. Gatlin
1975 "White flight and percent black: the tipping point re-examined." <i>Social Science Quarterly</i> 56:85-92. |
| Virginia | Amherst
Campbell
Dinwiddie
Henrico | Amherst
Campbell
Dinwiddie
Henrico | Lord, Dennis
1975 "School busing and white abandonment of public schools" <i>Southern Geographer</i> 15:81-92.
Lord, Dennis and John Catau
1977 "School desegregation policy and intra-school district migration." <i>Social Science Quarterly</i> 57:784-96. |
- REFERENCES
- Blalock, Herbert M.
1967 *Toward a Theory of Minority-Group Relations*. New York: Wiley.
- Bosco, James and Stanley Robin
1974 "White flight from court ordered busing?" *Urban Education* 9:87-98.
- Clotfelter, Charles T.
1976 "School desegregation, 'tipping' and private school enrollment." *Journal of Human Resources* 11:28-49.
- Coleman, James, Sarah D. Kelley and John Moore
1975 *Trends in School Segregation, 1968-1973*. Washington, D.C.: Urban Institute.
- Farley, Reynolds
1975 "School integration and white flight." Pp. 1-20 in Gary Orfield (ed.), *Symposium on School Desegregation and White Flight*. Washington, D.C.: Center for National Policy Review, Catholic Law School.
- Munford, Luther
1973 "White flight from desegregation in Mississippi." *Integrated Education* 11:12-26.
- Pettigrew, Thomas F. and Robert Green
1976 "School desegregation in large cities; a critique of the Coleman 'white flight' thesis." *Harvard Education Review* 46:1-53.
- Rossell, Christine
1976 "School desegregation and white flight." *Political Science Quarterly* 90:675-98.
- Stinchcombe, Arthur, Mary McDill and Dollie Walker
1969 "Is there a racial tipping point in changing schools?" *Journal of Social Issues* 25:127-36.
- Wegmann, Robert G.
1975 "Neighborhoods and schools in racial transition." *Growth and Change* 6:3-8.

MEASURES OF INEQUALITY *

PAUL D. ALLISON

Cornell University

American Sociological Review 1978, Vol. 43 (December):865-880

Although measures of inequality are increasingly used to compare nations, cities, and other social units, the properties of alternative measures have received little attention in the sociological literature. This paper considers both theoretical and methodological implications of several common measures of inequality. The Gini index is found to satisfy the basic criteria of scale invariance and the principle of transfers, but two other measures—the coefficient of variation and Theil's measure—are usually preferable. While none of these measures is strictly appropriate for interval-level data, valid comparisons can be made in special circumstances. The social welfare function is considered as an alternative approach for developing measures of inequality, and methods of estimation, testing, and decomposition are presented.

1. Defining Inequality

Although inequality has long been topic of intense interest to sociologists, few have bothered to carefully specify what they mean by the term. It is easy, of course, to distinguish perfect equality from a state of inequality. But given two different, unequal distributions of some social reward, how does one decide which distribution is the more unequal? The answer to this question would seem to be a prerequisite for any theory of the determinants and consequents of social inequality. Yet even Lenski's (1966) influential theory on the effects of economic surplus and democracy on inequality fails to include a definition of the dependent variable.

This lack of rigor created little difficulty so long as research on inequality emphasized the determinants of individual attainments. But recent efforts to test hypotheses explaining why some societies are less equal than others have necessitated the adoption of precise measures of inequality,

such as the Gini index or the standard deviation (Jackman, 1974; Robinson and Quinlan, 1977; Hewitt, 1977; Kelley and Klein, 1977). In the absence of clear criteria for choosing among the numerous measures of inequality, researchers have usually based their choice on convenience, familiarity, or on vague, methodological grounds. Nevertheless, the decision to rank one distribution as more unequal than another has theoretical as well as methodological implications. In fact, the choice of an inequality measure is properly regarded as a choice among alternative definitions of inequality rather than a choice among alternative ways of measuring a single theoretical construct.

This choice *can* make a difference. Although some have reported high correlations among different inequality measures (Alker and Russett, 1964), such correlations are relevant only for the particular populations and variables for which they are computed. And in one case of particular interest, it has been shown that the rank ordering of countries by income inequality can differ substantially with different measures of inequality (Atkinson, 1970; Yntema, 1933).

A major exception to the failure of sociological theorists to specify what they mean by greater and lesser inequality is the recent work of Peter Blau (1977a; 1977b). Blau assumes that inequality is a funda-

* Address all communications to: Paul D. Allison; SASS; 323 Uris Hall; Cornell University; Ithaca, NY 14853.

For helpful comments, I am indebted to Scott Feld, Lowell Hargens, J. Scott Long, Barbara Reskin and Alice Trisdorfer. This paper was written while I was supported by a postdoctoral fellowship from the Social Science Research Council.

mental characteristic of all graduated social parameters (quantitative status variables), and proposes that it be conceptualized as "the average difference in status between any two pairs relative to the average status" (1977b). Noting that the Gini index is an appropriate algebraic specification of this concept, he proceeds to consider in detail how changes in the distribution affect the Gini index (1977a:56-69).

Although Blau's formulation is certainly an improvement over previous sociological work, he makes a fundamental error in suggesting that inequality measures can be meaningfully applied to any quantitative variable. He also fails to consider recent developments in economic theory which have shed much light on the implications of alternative measures of inequality. Drawing on this literature, I shall discuss the characteristics and relative merits of some common measures of inequality. Section 2 introduces several measures of inequality and considers whether they satisfy basic theoretical criteria. Section 3 examines Blau's assumption that inequality measures can be meaningfully used with interval scales. The assumption is shown to be generally false, except under special conditions. Section 4 shows how various measures are related to the Lorenz curve. This sets the stage for section 5 which discusses the social welfare approach to inequality that is favored by some economists. Finally, section 6 reports basic results on estimation and hypothesis testing with sample data. Since much of the literature is concerned with income inequality, I will assume that income is the variable of interest. With a few notable exceptions, most of the results also can be applied to other quantitative variables.

2. Basic Criteria for Measures of Inequality

Suppose we have a population of n people, and each person receives annual income x_i with $i = 1, \dots, n$. It will be convenient for some applications to assume that the incomes are arranged in ascending order so that $x_1 \leq x_2 \leq \dots \leq x_n$. The problem is to get a single measure which characterizes every possible set of x_i 's in terms of inequality. What characteristics

should such a measure possess? At the very least, it should be zero when all individuals have identical incomes, and should have a positive value when two or more individuals differ. These conditions are satisfied by almost all the common measures of inequality including such familiar measures of dispersion as the range and the variance.

Many of these measures are ruled out by the criterion of *scale invariance*, which requires that multiplying everyone's income by a constant leaves the degree of inequality unchanged. This eliminates the variance, for example, since it quadruples when everyone's income is doubled. Although scale invariance is widely accepted as a desirable property of inequality measures, the justification is rarely made explicit. And since a few sociologists and economists have questioned this criterion,¹ it seems worthwhile to examine some of the arguments in its favor.

Since a change of units in which a variable is measured does not constitute any real change in the distribution of that variable, it seems appropriate to ask that a measure of inequality be invariant to such changes. Thus inequality of income should not depend on whether income is measured in dollars or yen. Obviously, scale invariant measures have this property. There is an enormous payoff in convenience since it becomes unnecessary to adjust for inflation or to deal with currency conversions.

It is less clear that *real* proportionate increases in everyone's income should leave inequality unchanged since, in absolute amounts, the rich benefit more than the poor. But note that increasing everyone's income by, say, 10% leaves the ratios of all pairs of incomes unchanged. Equivalently, each individual's proportionate share of the total national income remains the same. Hence, it is reasonable to say that the *relative* differences among individuals have not been altered. From a variety of perspectives, it is desirable that measures of inequality respond to relative

¹ Kelley and Klein (1977) and Kolm (1976) argue that proportional increases in income represent an increase in inequality. Taking the opposite point of view, Dalton (1920) and Sen (1973) suggest that inequality should *decrease* when all incomes are increased proportionately.

rather than to absolute differences (Blau, 1977a:57-9). For example, Easterlin (1974) presents evidence that individuals' self-reported happiness depends less on their absolute incomes than on their relative positions in the income distribution.

It may seem like a big jump to require that inequality be comparable across completely different sorts of quantities. Yet, it is common to hear assertions like "power is more unequally distributed than wealth in nation A." Without scale invariance, such assertions are meaningless. With scale invariance, we can, for example, compare the inequality of nations' energy consumption with the inequality of population size.

A final argument is that scale invariant measures (at least all those we shall consider here) respond in an intuitively appealing manner when a positive constant is *added* to everyone's score; specifically, they decline. Consider a three-person society with incomes of \$5,000, \$15,000, and \$25,000. Clearly, these differences would be of great importance to the individuals involved. If each were given a million dollars, however, the remaining differences would become trivial, and it would be reasonable to say that inequality had declined. Yet, the standard deviation would stay exactly the same. For all these reasons, the remainder of this paper will deal only with scale invariant measures.

Most measures of dispersion can be converted into scale invariant measures of inequality by dividing by the mean or some function of the mean. The *coefficient of variation* V , for instance, is just the standard deviation divided by the mean:

$$V = \frac{\sigma}{\mu}. \quad (1)$$

Similarly, the *relative mean deviation*, also known as Schutz's (1951) coefficient, is defined as

$$D = \frac{\frac{1}{n} \sum_{i=1}^n |x_i - \mu|}{2\mu}, \quad (2)$$

where $|\cdot|$ is the absolute value function.

Perhaps the most commonly used measure of inequality is the *Gini index*, G , which is usually defined in terms of the

Lorenz curve (see section 4). However, the following equivalent definition makes it clear that the Gini index is a measure of dispersion divided by twice the mean:

$$G = \frac{\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{2\mu}. \quad (3)$$

The numerator of this expression, known as Gini's coefficient of mean difference (Kendall and Stuart, 1977: 48), is the average absolute difference between all pairs of individuals. An equivalent formula (Dasgupta et al., 1973) is more mathematically tractable and computationally convenient for individual-level data:

$$G = \frac{2}{\mu n^2} \sum_{i=1}^n ix_i - \frac{n+1}{n}. \quad (4)$$

Notice that the first term in (4) involves a weighted sum of all the scores, where the weight applied to each score is its rank in the distribution.

Based on information theory, Theil (1967: 92) proposed the measure

$$T = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i}{\mu} \right) \log \left(\frac{x_i}{\mu} \right). \quad (5)$$

(Natural logarithms are used throughout this paper.) Simple algebra reduces this to a formula that is more computationally convenient and also reveals that T is a measure of dispersion divided by the mean:

$$T = \frac{\frac{1}{n} \sum_{i=1}^n x_i \log x_i - \mu \log \mu}{\mu}. \quad (6)$$

When $x_i = 0$, $x_i \log x_i$ is also defined to be 0.

One scale invariant measure that cannot be expressed as a ratio of a measure of dispersion to the mean is the *variance of the logarithms* L . This is obtained by taking the logarithm of each income and computing the variance of the transformed scores. Thus, if $z_i = \log x_i$ for all i ,

$$L = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2. \quad (7)$$



L is undefined when the distribution includes incomes of zero. Although many other measures of inequality have been proposed (Alker and Russett, 1964; Martin and Gray, 1971; Ray and Singer, 1973), they either fail to satisfy the preceding criteria, have very restricted applications, or are simple monotone functions of those defined here.

How is one to choose among these five scale-invariant measures? Considerable headway can be made by applying the *principle of transfers*. Dalton (1920) argued that measures of inequality ought to increase whenever income is transferred from a poorer person to a richer person, regardless of how poor or how rich or the amount of income transferred. Not only does this principle have considerable intuitive appeal but, as we shall see, it also has important relationships with the Lorenz curve and the social welfare approach to inequality measurement.

Two of the scale-invariant measures fail to satisfy the principle of transfers (Atkinson, 1970). The relative mean deviation is not affected by transfers between persons who are both below the mean or both above it, and it will therefore be dropped from further consideration. The variance of the logarithms responds appropriately to transfers at lower income levels. But at high income levels (greater than 2.718 times the geometric mean), it actually decreases with a transfer from a (relatively) poorer to a richer individual.² Although this is a serious limitation, the variance of the logarithms has very desirable inferential properties which will be discussed in section 6.

Sensitivity to transfers. Among those measures of inequality which satisfy the principle of transfers, there are important differences in sensitivity to transfers at different points on the scale (Atkinson, 1970). Suppose we transfer h dollars from a person with income x_i to another person with income x_j , where $x_i \leq x_j$. All other incomes remain the same. Let V_1 and V_2 be the coefficients of variation before and after the

transfer, respectively. It can be shown (Dalton, 1920) that

$$V_2^2 - V_1^2 = c h(x_j - x_i) + c h^2, \quad (8)$$

where c is positive and depends only on the mean and the number of observations (the same is true for c' and c'' used below). This result says that V is equally sensitive to transfers at all income levels. Thus, a transfer of \$100 from a person earning \$5,000 to another earning \$6,000 has the same impact as a transfer of \$100 from a person earning \$50,000 to another earning \$51,000.

The Gini index is peculiar in that its sensitivity to transfers depends on individuals' ranks rather than their numeric scores. Using formula (4), it is easily shown that, for a transfer of h from x_i to x_j ,

$$G_2 - G_1 = c'h(j - i), \quad (9)$$

where G_2 and G_1 are the values of the Gini index before and after the transfer, and i and j are the ranks of incomes x_i and x_j . This result says that the change in G depends on the number of individuals with incomes lower than x_j and higher than x_i . In the U.S. today, there are many more persons in the \$10,000–\$11,000 interval than there are in the \$50,000–\$51,000 interval. It follows that a transfer from one person earning \$10,000 to another earning \$11,000 will have more effect on G than an equal transfer from a person earning \$50,000 to another earning \$51,000. But there are also fewer people in the \$4,000–\$5,000 interval than in the \$10,000–\$11,000 interval. Thus, for a typically shaped income distribution, the Gini tends to be most sensitive to transfers around the middle of the distribution and least sensitive to transfers among the very rich or the very poor.

To get a simple expression for the effect of a transfer on Theil's measure T , it is necessary to use a limiting argument. Let ΔT be the change in T resulting from a transfer of h from x_i to x_j . As h goes to 0, the limiting expression for ΔT can be shown to be

$$\Delta T = c''h \log(x_j/x_i). \quad (10)$$

Whereas the change in V depended on the differences between the incomes, the change

² Creedy (1977) argues that the extent to which the variance of the logarithms violates the principle of transfers is very minor for most empirical distributions.

in T depends on the ratio of the incomes. As a consequence, transferring \$100 from a person earning \$5,000 to one earning \$6,000 has approximately the same effect on T as a transfer of \$100 from a person earning \$50,000 to another earning \$60,000. This change in T is approximately nine times as large as the change resulting from a transfer of \$100 from a person earning \$50,000 to another earning \$51,000. The lower the level of income, the more sensitive T is to transfers. (The variance of the logarithms is similar in this respect.)

Do these differences in sensitivity give any basis for choosing among G , V and T ? If it is assumed that income has diminishing marginal utility, then a transfer of income among low income earners would be more consequential (for them) than a transfer of an equal amount among high earners. Since Theil's index T reflects such a difference, it would seem to have the advantage. On the other hand, when the variable of interest does not have diminishing marginal utility or when its utility or value (if any) is irrelevant to the analysis, then the "flat" response of the coefficient of variation might be desirable. Thus, depending on the context, V might be a good measure of inequality of age, city size, years of schooling, etc. The sensitivity of the Gini index depends on the shape of the frequency distribution, and it is not easy to think of cases where this property would be desirable in itself. Of course, since many distributions tend to be somewhat bell-shaped, this means that G tends to be most sensitive in the middle range. If one is most concerned about changes in inequality among middle income earners, then G might be a good choice.³

These differences in sensitivity can have important consequences. In comparing income inequality across nations, Atkinson

(1970) concluded that measures which were most sensitive in the lower range of incomes tended to show relatively less inequality in developing nations and more inequality in developed nations. On the other hand, measures which were sensitive in the higher ranges tended to favor the developed nations. He suggested that this was due to the tendency for developing nations to have a large, homogeneous population of poor together with great inequality among the rich (cf. Blau, 1977a: 9).

Upper and lower bounds. Sometimes inequality measures are chosen on the basis of their upper and lower bounds, or how they respond to changes in the population size. In infinite populations the Gini index varies between 0 and 1 while the coefficient of variation and Theil's measure vary between 0 and infinity. These bounds should imply no preference, however, since simple (nonunique) transformations can produce any desired bounds. To make the coefficient of variation have an upper bound of 1, simply take $V/(V+1)$. To make the Gini index vary between minus and plus infinity, take its logit: $\log(G/(1-G))$. The logit transformation has some merit when the Gini index is to be used as a dependent variable in a regression analysis since it avoids the usual problems associated with bounded dependent variables (Nerlove and Press, 1973).

In finite populations or samples, the Gini index has an upper bound of $1-1/n$, the coefficient of variation has an upper bound of $\sqrt{n-1}$ and Theil's measure has an upper bound of $\log n$. For all three measures, the upper bound is reached when one individual has everything and everyone else has nothing. Theil (1967:92) has argued that this dependence on n is desirable since a two-person society in which one person has everything is, intuitively, less unequal than a million-person society in which one person has everything. Nevertheless, there may be some situations in which it is desirable to have a measure whose upper bound does not depend on n (Ray and Singer, 1973).⁴ Again, however,

³ Using simulated distributions, Champenowne (1974) reached very similar conclusions about these differing sensitivities. He found that the coefficient of variation was most sensitive to inequality of extreme wealth, the Gini index was most sensitive in the middle range, and the variance was most sensitive to inequality of extreme poverty. Contrary to the results here, however, he found that Theil's index behaved more like V than like L .

⁴ Ray and Singer (1973) proposed a new index CON whose range did not depend on n . It can be shown, however, that CON is just the coefficient of variation divided by its upper bound.

it is a simple matter to divide V , G and T by their upper bounds to obtain measures that vary between 0 and 1 independently of n (Martin and Gray, 1971). Alternatively, it may be preferable to use $V^* = V/(\sqrt{n-1} - V)$ and $T^* = T/(\log n - T)$ to get measures which vary between 0 and infinity.

3. Inequality and Interval Scales

Discussions of the coefficient of variation and the Gini index in statistics texts frequently end with a warning that the measures are only appropriate for variables measured on a ratio scale, like income or age, which have a theoretically fixed zero point (Mueller et al., 1977). As Kendall and Stuart (1977:48) put it:

The coefficients both suffer from the disadvantage of being very much affected by . . . the value of the mean measured from some arbitrary origin, and are not usually employed unless there is a natural origin of measurement or comparisons are being made between distributions with similar origins.

Jencks et al. (1972:352) have noted that this tends to rule out the application of inequality measures to such common variables as IQ scores and occupational prestige which are, at best, measured on interval scales.

Blau (1977a; 1977b), choosing to disregard this caveat, however, claims that "the Gini index is substantively appropriate for any status criterion" including mathematical aptitudes and intelligence. The aim of this section is to show that Blau is only partly right. In general, one cannot make valid inferences about inequality using interval-level data. But under certain conditions, some of which can be easily checked, such inferences are valid.

The presentation can be simplified by first observing a special relationship between V and G . Kendall and Stuart (1977: 41) show that the variance can be expressed as one-half the average squared difference between all pairs of individuals:

$$\sigma^2 = \frac{1}{2n^2} \sum_{i=1}^n \sum_{j=1}^n (x_i - x_j)^2. \quad (11)$$

Using this fact with formulas (1) and (3), we can define a general family of inequality measures that includes both the Gini index and the coefficient of variation:

$$I = \frac{\left(\frac{1}{2n^2} \sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|^r \right)^{\frac{1}{r}}}{\mu}. \quad (12)$$

When $r = 1$, I is the Gini index. When $r = 2$, I becomes the coefficient of variation.⁵

This result enables us to consider some common properties of G and V by referring only to I . For example, I is clearly scale invariant, which can be expressed as $I(cX) = I(X)$ where c is any constant.⁶ If X is an interval-scale variable, its origin (zero point) may be arbitrarily changed by adding a constant to every score. The measure I is quite sensitive to changes in origin, however; it can be readily verified that

$$I(X + c) = \left(\frac{\mu}{\mu + c} \right) I(X). \quad (13)$$

By appropriate choice of c , one can thus make $I(X + c)$ take on any real value, a quite undesirable result.⁷ Clearly these measures of inequality are useless for characterizing a single distribution if measurements are only at the interval level.

What about comparisons of inequality for two distributions, each measured from the same, arbitrary origin? Could we not, for example, compare inequality of Duncan SEI scores for the occupations of blacks and whites? Although Kendall and

⁵ It can be shown that I satisfies the principle of transfers for all $r \geq 1$. For another family of inequality measures which includes the coefficient of variation and Theil's measure, but not the Gini index, see Gastwirth (1975).

⁶ In this notation, X is an $n \times 1$ vector of scores. Depending on the context, c is either a scalar or an $n \times 1$ vector of equal constants.

⁷ The usual (0, 1) bounds on the Gini index apply only if all scores are nonnegative. Formulas (3) and (4) make it clear that G can be readily computed even if some or all scores are negative. In that case, however, the usual interpretation in terms of Lorenz curves is invalid (see section 4). Equation (13) also holds for D , the relative mean deviation, even though D is not a special case of (12). It also can be shown that T and L decrease when a positive constant is added to all scores and, hence, these measures also have arbitrary values when applied to interval scales.

Stuart suggest that such comparisons would be meaningful, it is easy to construct counterexamples. Suppose we wanted to compare the inequality of two distributions X and Y , both measured on the same interval scale. We observe $I(X) = .50$, $I(Y) = .60$, $\mu_X = 20$, and $\mu_Y = 10$. Let us change the origin of measurement by adding 10 to each score. Using formula (13), we get $I(X + 10) = .33$ and $I(Y + 10) = .30$. Thus, different origins lead to different conclusions about the relative inequality of the two distributions.

An empirical example is provided by Blau's (1977a: 211) reanalysis of data reported by Hauser et al. (1975) who give the frequency distribution among 12 occupational categories of U.S. men at selected ages in 1952, 1962, and 1972. Using the Duncan SEI scores corresponding to the occupational categories, Blau calculated the Gini index for men aged 35-44 and noted the trend in inequality. The Gini values for the three periods were .353, .300, and .318—a clear decline in inequality over time. The corresponding mean scores were 31.8, 37.5, and 40.4. But since SEI scores are only interval level at best, we can freely change the origin by adding an arbitrary constant to all scores. A constant of 50 yields Gini values of .137, .141, and .142, a lower level of inequality but increasing rather than decreasing with time. Blau's prescription is clearly unacceptable.

Nevertheless, there are circumstances in which valid comparisons of inequality can be made for two distributions measured on the same interval scale. These cases arise when we know (or believe) that there is a nonnegative ratio scale underlying the interval scale that we observe. For instance, if we wish to measure social power, an interval scale may be the best that we can obtain. Yet zero power or powerlessness is clearly a meaningful concept and negative power is not. It is therefore reasonable to believe that there is a true ratio scale underlying our measurements of power. This is important because it puts constraints on the constant that can be added to or subtracted from the interval measure.

More generally, suppose X^* is a nonnegative ratio-level variable and X is an inter-

val-level variable obtained by adding an unknown constant to X^* , i.e., $X = X^* + c$. If we observe a set of X 's, we can be sure that c is no greater than the smallest value of X . Otherwise X^* would have negative values which are disallowed by assumption. Now suppose that we also have $Y = Y^* + c$, where Y^* is another nonnegative ratio-level variable measured on the same scale and origin as X^* . It follows that X and Y also have a common scale and origin. We observe $I(X) > I(Y)$, as well as the means μ_X and μ_Y . Under what circumstances can we conclude that $I(X^*) > I(Y^*)$? The answer is that $I(X^*) > I(Y^*)$ whenever there is no permissible value of c such that $I(X^*) \leq I(Y^*)$.

The following test for this condition is proved in the Appendix. First, make sure that the origin of X and Y is such that all values are nonnegative. Then calculate

$$c^* = \frac{\mu_X \mu_Y [I(X) - I(Y)]}{\mu_X I(X) - \mu_Y I(Y)}. \quad (14)$$

If c^* is greater than the minimum value of either X or Y , we can conclude that $I(X^*) > I(Y^*)$. If c^* is less than or equal to the minimum values of both X and Y , no conclusion may be drawn. When $\mu_X I(X) = \mu_Y I(Y)$, c^* is undefined, but this also implies that $I(X^*) > I(Y^*)$. If the minimum values of X and Y are not readily available, then $c^* > \mu_X$ or $c^* > \mu_Y$ obviously insures that it is greater than one of the minimum values.

As an example, I apply this test to data on occupational status of blacks and whites reported by Farley (1977). In 1959, black males had a mean occupational status of 18.9 and a coefficient of variation of .815; in the same year, white males had a mean of 37.2 and a coefficient of variation of .608. The value of c^* , computed from formula (14), is -20.17 . Since this is necessarily less than the minimum value for both blacks and whites, no conclusion about inequality is possible. In 1969, black females had a coefficient of variation of .773 and a mean of 28.6; white females had a coefficient of variation of .491 and a mean of 42.6. This yields $c^* = 288.43$ which is greater than the mean of either group. We conclude that in this sample black females were more unequal in occu-

pational status than white females. This presumes, of course, that there is an underlying nonnegative ratio scale of occupational status, an assumption some may wish to question. It also says nothing about possible sampling errors.

4. Measures of Inequality and the Lorenz Curve

The Gini index, as noted earlier, is usually defined in terms of the Lorenz curve (e.g., Ray and Singer, 1973). It so happens that all scale-invariant measures of inequality which satisfy the principle of transfers have a simple relationship to the Lorenz curve. This fact makes it possible to formulate a criterion for greater and lesser inequality which transcends the several measures we have been considering. It is also closely bound up with the social welfare approach to inequality which will be discussed in the next section.

Suppose we rank order all the individuals in a population by income, from lowest to highest. For each rank, we may calculate the proportion of the population at that rank or below, and also the proportion of the total income that is earned by people at that rank or below. We may find, for example, that the poorest .25 of the population earns only .05 of the nation's income. If we plot the relationship between these two proportions for every rank, we get the Lorenz curve. Three such curves are shown in Figure 1. Line A, the diagonal straight line, is the Lorenz curve under the condition of perfect equality. It is often taken as a reference point for other observed curves. If any persons have unequal incomes, the Lorenz curve will fall below the line of perfect equality, indicating that the poorest y percent of the population earns less than y percent of the income for some value(s) of y .

The Gini index is equal to twice the area between the Lorenz curve and the line of perfect equality. Clearly line C, which falls entirely below line B, corresponds to a larger Gini index than line B. As many authors have shown (Morris, 1972), this conclusion also applies to any scale-invariant measure which satisfies the principle of transfers. Let us state this more formally.

Let H be a scale-invariant measure of inequality that satisfies the principle of transfers, and let X and Y be two distributions. If the Lorenz curve for Y is never above and is somewhere below the Lorenz curve for X , then $H(X) < H(Y)$. The upshot is that whenever one Lorenz curve "dominates" another, in this sense, it makes little difference whether one uses the Gini index, the coefficient of variation or Theil's measure. All three will give the same rank ordering. It is not uncommon, however, for two Lorenz curves to intersect. In such cases, the Gini index may give one rank ordering while the coefficient of variation gives another.

5. Inequality and Social Welfare

In 1920, Dalton argued that the choice of an inequality measure typically involves an implicit normative judgment as to whether one distribution of income is to be preferred, in some sense, to another. He concluded that we would be in a better position to devise and choose measures of inequality if we could make these normative criteria more explicit. The problem may be formulated as follows: Suppose that we have a fixed total income that is to be distributed among n persons. We assume that for any possible distribution of incomes there is a number W indicating the desirability of that distribution. In functional notation

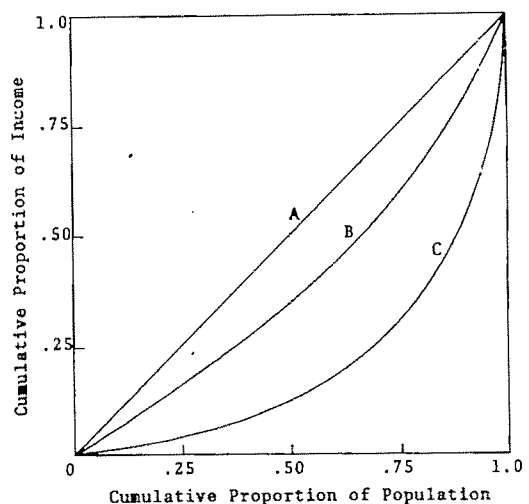


Figure 1. Lorenz Curves for Three Distributions of Income

$$W = W(x_1, x_2, \dots, x_n) = W(X). \quad (15)$$

W is called the social welfare function. If we can specify the form of W , and if W reaches a maximum when all incomes are equal, then it is plausible to take a decreasing function of W as a measure of inequality.

This approach was largely ignored until Aigner and Heins (1967) proposed some alternative functions for W and derived corresponding measures of inequality. Atkinson (1970) greatly increased interest in the approach by demonstrating that plausible constraints on W implied an important relationship between W , the Lorenz curve, and the principle of transfers. He began by assuming that

$$W = \sum_{i=1}^n U(x_i). \quad (16)$$

$U(x_i)$ may be thought of as the utility of income x for individual i . Hence, the total social welfare is just the sum of the individual utilities. He further assumed that all individuals have the same utility function and that $U(x_i)$ is concave, i.e., there is diminishing marginal utility from increasing income. These last two assumptions imply that W will be maximized when all incomes are equal. Atkinson's major contribution was to show that, under these conditions, an ordering of the Lorenz curves implies an ordering of social welfare. Specifically, if we are given two distributions X and Y and the Lorenz curve for X is never below and somewhere above the Lorenz curve for Y , then $W(X) > W(Y)$. This result can also be expressed in terms of the principle of transfers. If a distribution Y can be obtained from a distribution X by a sequence of transfers from poorer to richer individuals, then $W(X) > W(Y)$.

Atkinson's conditions on W may seem rather restrictive, but they have since been greatly generalized by Dasgupta et al. (1973), Sen (1973), and Rothschild and Stiglitz (1973). The generalized versions of the theorem do not require that W be a sum of individual utility functions. What is necessary is that W be symmetric, continuous, monotonic, and "locally equality preferring." For an especially lucid account

of these generalizations, see Rothschild and Stiglitz (1973).

The force of these theorems is to strengthen the theoretical foundation of the Lorenz curve and the measures of inequality that are associated with it. Given a fixed total income, if the Lorenz curve for one distribution X lies above the Lorenz curve for another distribution Y , we can be sure that X is preferable to Y under a broad class of functions defining preferability. Fixing the total income may seem like a restrictive condition, but it should not be a surprising one. Indeed, we would hardly expect that a perfectly equal distribution of \$10 among ten persons would produce greater social welfare than a slightly unequal distribution of \$1,000 among the same ten persons. Nevertheless, Rothschild and Stiglitz (1973) give a generalized definition of Lorenz dominance that allows for differences in total income in two distributions.

Let us return to Dalton's original aim of constructing a measure of inequality based on the social welfare function. The problem, of course, is that the choice of a social welfare function is normative, not empirical, and it is difficult to achieve any agreement on what that function should be. Perhaps the most widely acceptable class of functions is the additive, concave welfare function defined in (16). For this class, Atkinson (1970) showed that the following inequality measure was especially appropriate:

$$A = 1 - \frac{n}{1} \left[\sum_{i=1}^n \left(\frac{x_i}{\mu} \right)^{1-e} \right]^{\frac{1}{1-e}}, \quad (17)$$

where $e > 0$. The parameter e determines the "inequality aversiveness" of the measure. As e rises, A becomes more sensitive to transfers among lower incomes and less sensitive to transfers among top income recipients. In the limit, as $e \rightarrow 1$, A goes to $1 - M/\mu$ where M is the geometric mean and μ is the arithmetic mean. A is also scale invariant and satisfies the principle of transfers.

The advantage of A is that it provides a flexible, theoretically-based approach to the sensitivity question discussed in section 2. Thus, one may choose e to conform to

one's judgment about what portions of the distribution are most relevant to the analysis. For some empirical applications of A, see Atkinson (1970), Bartels and Nijkamp (1976), and Williamson (1977). In all three cases, the authors replicate the analysis using several different values of e .

6. Estimation and Testing

To this point I have dealt only with measures of inequality applied to entire populations. In this section, I consider several issues related to estimation and testing with sample data, both grouped and ungrouped. As a prelude to the case of grouped data, I also present formulas for decomposing inequality into within- and between-groups components.

Ungrouped data. The usual approach to estimating measures of inequality is to apply the population formulas (1) through (7) to the sample in hand. While this produces consistent estimators in most cases, the resulting sampling distributions tend to be quite complicated, making it difficult to obtain standard errors, confidence intervals and test statistics. There is also the possibility that more efficient estimators might be available.

An alternative approach is that of maximum likelihood, which produces estimators that are approximately unbiased and efficient in large samples. To derive maximum likelihood estimators (MLEs), it is necessary to assume that the data come from a known family of distributions. While most statistical theory is built upon the normal distribution, this is an unsuitable family for the estimation of measures of inequality. The normal distribution is symmetrical and ranges from minus infinity to plus infinity. By contrast, measures of inequality are usually used only for variables which are inherently nonnegative, and the distributions are often positively skewed. The *lognormal* distribution has both these characteristics, and also has been found to provide a reasonably good fit to frequency distributions for a wide variety of positively-valued variables in several fields (Aitchison and Brown, 1957).

The lognormal distribution can be simply defined as follows. Suppose we have

two random variables X and Y , and $Y = \log X$. If Y is normally distributed with mean R and variance L , then X is said to have a lognormal distribution with parameters R and L . Notice that L is just the variance of the logarithms which we considered in section 2. The fact that it is also the variance of a normally distributed random variable means that we can apply the usual normal theory for inferences about variances.

For the lognormal distribution, most of the inequality measures defined earlier can be expressed as increasing functions of L (Aitchison and Brown, 1957; Theil, 1967: 97). Thus

$$\begin{aligned} T &= L/2; \\ V &= (e^L - 1)^{1/2}; \\ G &= 2 \Phi(\sqrt{L/2}) - 1; \end{aligned} \quad (18)$$

where $\Phi(\cdot)$ is the cumulative distribution function for a standard normal variable. That is, $\Phi(a)$ is the probability that a normally distributed variable, with a mean of zero and a variance of one, is less than or equal to a .

To get MLEs of all these measures, we apply a basic principle of likelihood estimation (Cox and Hinkley, 1974: 302). Suppose we have a parameter A which can be expressed as a function of another parameter B , i.e., $A = f(B)$. Then if \hat{B} is the MLE of B , it follows that $\hat{A} = f(\hat{B})$ is the MLE of A . This means that if we have the MLE of L , we can readily get MLEs of T , V , and G by substituting into (18). If the data consist of a simple random sample of n observations, it is well-known that the sample variance of $\log x_i$ (given in (7)) is the MLE of L .

As a numerical example, suppose that for a sample of 101 persons, the variance of the logarithm of income is .84. Then using (18) we get estimates of $\hat{T} = .42$, $\hat{V} = 1.15$, and $\hat{G} = .48$. In computing G , one can use the normal probability table found in almost any basic statistics text.

The asymptotic standard errors of \hat{L} , \hat{T} , \hat{V} and \hat{G} are easily derived (Kendall and Stuart, 1977: 258, 247), but it is not ad-

visible to use them in constructing confidence intervals unless the sample is quite large. A better approach is to use the usual formula for constructing confidence intervals around the variance of a normal variate. A $(1 - \alpha)100$ percent confidence interval around L is given by

$$\frac{n\hat{L}}{\chi^2(n-1, \frac{\alpha}{2})} \leq L \leq \frac{n\hat{L}}{\chi^2(n-1, \frac{1-\alpha}{2})}, \quad (19)$$

where $\chi^2(n-1, b)$ is the value in a chi-square distribution with $n-1$ degrees of freedom which cuts off the upper b of sample values (Hays, 1963: 345).⁸ In the example just given, a 95% confidence interval around L is $.65 \leq L \leq 1.13$. To obtain confidence interval for T , V , and G , simply substitute the upper and lower bounds for L into (18) which yields the corresponding upper and lower bounds for the derived inequality measures. In the example, this produces 95% confidence intervals of $.33 \leq T \leq .57$, $.96 \leq V \leq 1.45$, and $.43 \leq G \leq .64$.

It is also easy to test whether the level of inequality is the same in two distributions. Suppose we have two independent random samples of n_1 and n_2 observations from two different lognormal distributions. Since each of the four inequality measures is an increasing function of any of the others, $L_1 = L_2$ implies that $T_1 = T_2$, $V_1 = V_2$, and $G_1 = G_2$. Thus, we need only test the null hypothesis that $L_1 = L_2$. The test statistic is just \hat{L}_1/\hat{L}_2 which has an F distribution with n_1-1 and n_2-1 degrees of freedom under the null hypothesis (Hays, 1963: 351).

Although the assumption of a lognormal distribution provides an elegant and easy-to-use solution to the problem of inferences about measures of inequality, it is not without its limitations. First, it essentially does away with the differences in the various inequality measures discussed in section 2. Since any of the measures will give the same rank order for a set of lognormal distributions, one might just as well use L

alone. Converting to T , G , or V by the formulas in (18) is only useful to achieve comparability with other studies. This is surely an advantage when the data truly come from a lognormal distribution. But departures from lognormality expose one to the peculiarities of L which were discussed in section 2. Second, neither the confidence interval in (19) nor the F test for comparing two populations is very robust to departures from lognormality (Hays, 1963: 352). Only very large samples can compensate for serious departures.⁹ Finally, the methods fail when the sample includes values of zero since $\log 0$ is undefined and, hence, L is undefined.

Decomposition. It is often desirable to decompose the inequality in a population into inequality between groups and inequality within groups. For example, it would be of considerable interest to decompose the inequality of household income in the U.S. into inequality between states and inequality within states. Theil (1967: 123-7) has given several decomposition formulas and I shall reproduce some of them here using slightly different notations.

Suppose the population can be divided into J mutually exclusive and exhaustive groups. For each group $j=1, \dots, J$ we know \bar{X}_j , the arithmetic mean income in group j , p_j , the proportion of the population in group j , M_j , the geometric mean income, and T_j , V_j , and L_j , the level of inequality in each group for three different measures. The decomposition of Theil's index for the total population is

$$T = \sum_{j=1}^J \left(\frac{p_j \bar{X}_j}{\bar{X}} \right) \log \left(\frac{\bar{X}_j}{\bar{X}} \right)$$

⁹ A plausible alternative to the lognormal distribution is the two-parameter gamma family, which also has an origin at zero and is positively skewed. Since one of the parameters α determines G , V , and T , one can take the same approach used here for the lognormal: first get the MLE of α and use that to get MLEs of G , V , and T . Estimation is considerably more difficult, however, because it requires the iterative solution of an equation involving a digamma function. Moreover, the equations relating α to G and T involve digamma and incomplete beta functions which are inconvenient to evaluate (Salem and Mount, 1974).

⁸ For large n , it may be convenient to use the normal approximation to the chi-square distribution. See, e.g., Hays (1963:347).

$$+ \sum_{j=1}^J \left(\frac{p_j \bar{X}_j}{\bar{X}} \right) T_j, \quad (20)$$

where $\bar{X} = \sum_{j=1}^J p_j \bar{X}_j$ is the grand mean.

The first term on the right hand side is the between-groups component. It is equivalent to the value of T that would be obtained if everyone in each group received the mean income for that group. The second term on the right hand side is a weighted average of the within-group values of T . The weight for group j is the fraction of the total income that is earned by group j . For an extensive empirical application of (20), see Theil (1967: 98-114).

For L , the variance of the logarithms, we have

$$L = \sum_{j=1}^J p_j (\log M_j - R)^2 + \sum_{j=1}^J p_j L_j, \quad (21)$$

where $R = \sum_{j=1}^J p_j \log M_j$. The first compo-

nent is the value of L that would be obtained if everyone in group j received income M_j . The second component is again a weighted average of the within-group values of L , but here the weights are just the fractions of the population contained in each group.

Finally, we have

$$V^2 = \sum_{j=1}^J \frac{p_j (\bar{X}_j - \bar{X})^2}{\bar{X}^2} + \sum_{j=1}^J \left(\frac{p_j \bar{X}_j}{\bar{X}^2} \right) V_j^2. \quad (22)$$

Notice that this is a decomposition of the *squared* coefficient of variation. The interpretation is similar to that for Thiel's measure, with one exception: the weights for V_j^2 do not have an obvious interpretation and they do not sum to one. Instead, they sum to one plus the between-groups component, a somewhat undesirable result.

The Gini index cannot be conveniently decomposed.¹⁰

Grouped data. Frequently the researcher will wish to estimate the inequality in some population when individual-level data are unavailable. Either the data are grouped on some exogenous criterion, e.g., geographic units, or the grouping is by interval of the variable of interest. Let us first consider the exogenous criterion.

Suppose we wish to estimate the inequality of household income in a state, but the data are grouped by census tract. In each tract, we know the mean income \bar{X}_j and p_j , the proportion of households in tract j . In order to get a lower bound on T and V , the obvious approach is to compute the between-groups components in (20) and (22). That this is a lower bound is evident from the fact that the true values of T and V must also include the within-groups components which could conceivably be quite large.

The geometric means will usually not be available to compute the between-groups components of L , the variance of the logarithms. However, we can still compute the variance of the logarithms that would be obtained if everyone in each group received the arithmetic mean income for that group. This is given by

$$\sum_{j=1}^J p_j (\log \bar{X}_j - R')^2, \quad (23)$$

where $R' = \sum_{j=1}^J p_j \log \bar{X}_j$. Similarly, we can

compute the Gini index that would be obtained if everyone in each group received the arithmetic mean income for that group. A convenient formula given by Blau (1977b) is

$$\frac{\sum_{j=1}^J \bar{X}_j p_j (q_j - r_j)}{\bar{X}}, \quad (24)$$

where q_j is the proportion that are in

¹⁰ Pyatt (1976) attempts a decomposition of the Gini index, but ends up with three components instead of two.

groups with means less than \bar{X}_j and r_j is the proportion in groups with means greater than \bar{X}_j . Hence $p_j + q_j + r_j = 1$ for all j . As in the previous cases, this will give a lower bound on the true value of G .

The other sort of grouped data consists of a grouped frequency distribution. We are given a set of cutpoints a_j for $j=1, \dots, J+1$, and p_j the proportion of observations falling in the interval $[a_j, a_{j+1})$ for $j=1, \dots, J$. In most cases, a_1 will be zero, and in some cases a_{j+1} will be unspecified. In some cases, we also know \bar{X}_j , the mean in each interval. When \bar{X}_j is not available, most methods require an estimate, such as the arithmetic midpoint, $\frac{1}{2}(a_{j+1} + a_j)$. The

most common method for estimating measures of inequality from grouped frequency distributions is to treat the data as if the grouping were based on some exogenous criterion. That is, the inequality measures are computed as though each person's score were the interval mean or midpoint. This is a lower bound, as we have just seen. When the data are grouped by intervals, however, it is also possible to compute sharp upper bounds on V , T and G by methods recently introduced by Gastwirth (1972; 1975) and Mehran (1975). An alternative approach begins with the assumption that the data are drawn from an underlying probability distribution. Aitchison and Brown (1957: 51) give the MLE of L under the assumption of a lognormal distribution. Aigner and Goldberger (1970) assume a Pareto distribution and give several methods for estimating the parameter α , which also determines many of the inequality measures. Kakwani and Podder (1973; 1976) specify the distribution by specifying the Lorenz curve; they then obtain efficient estimators for all inequality measures that depend on the Lorenz curve.

Conclusion

One approach to defining inequality is to choose one of the common measures of inequality. The range of choice is narrowed considerably by requiring that (a) inequality be invariant to proportionate increases

or decreases in everyone's score, and (b) any transfer from an individual with a lower score to another individual with a higher score represents an increase in inequality. This leaves three measures: the Gini index, the coefficient of variation and Theil's measure.

Although the Gini index is the most popular measure of inequality, it does not appear to possess any special advantages over the other two measures. Moreover, it is the most difficult to compute and cannot be decomposed into inequality within and between groups. Because its sensitivity to transfers decreases as scores increase, Theil's measure is especially desirable for measuring inequality of income, or other social rewards having diminishing marginal utility. For variables like age, where utility is neither strictly increasing nor especially relevant, the flat sensitivity of the coefficient of variation makes it the appropriate choice. The fact that it is also easily obtained from standard computer output makes it particularly useful for exploratory work.

None of these inequality measures is appropriate for interval-level variables which lack a theoretically fixed zero point. If they are applied to such data, the numerical values are essentially arbitrary. Nevertheless, we saw in section 3 that there are special situations in which valid comparisons of inequality can be made for two interval-level distributions. These situations arise when it can be assumed that there is an underlying nonnegative ratio scale. If such an assumption cannot be made, the only recourse is to use a scale-dependent measure of dispersion such as the standard deviation.

It is a simple matter to get efficient estimates of the Gini index, the coefficient of variation, and Theil's measure under the assumption that the data come from a lognormal distribution. Simply take the logarithm of each score and calculate the variance of these transformed scores. Using formulas in section 6, we can then easily obtain maximum likelihood estimates of the three measures. Methods for confidence intervals, hypothesis tests, and grouped data also were presented.

Many economists have taken the position that instead of choosing one of the traditional inequality measures, one should first specify a social welfare function and then derive an appropriate measure of inequality. Unfortunately, there is little consensus on the general form of the social welfare function, let alone the parameter values for any particular functional form. Atkinson (1970) also has warned that a social welfare function for the distribution of income among persons may be entirely inappropriate for the social welfare of industrial concentration, another area in which measures of inequality are frequently employed (Hall and Tideman, 1967). The same may be said for such social rewards as power and prestige.

Sen (1973) has suggested that current notions of inequality may be too imprecise for any conventional measure of inequality. In his view, the best we may be able to achieve without exceeding our understanding is a partial ordering: for some pairs of distributions we can say that one is more unequal than the other, while for other pairs we suspend judgment. Such a partial ordering is produced by the criterion of Lorenz dominance discussed in section 4. If one Lorenz curve lies nowhere below and somewhere above another, the first distribution is the more equal. For distributions whose Lorenz curves cross, no ranking is possible. Although Sen's position is extreme, it is worthwhile keeping in mind that the Lorenz dominance criterion is virtually unquestioned. Thus, if a particular hypothesis about inequality can be demonstrated by Lorenz dominance, confidence in the result is greatly strengthened.

APPENDIX

Proof of Test for Differences in Inequality When Variables Are Measured on an Interval Scale

Let X be an $n \times 1$ vector of observations on an interval scale and let Y be an $m \times 1$ vector of observations on the same interval scale. Let X^* and Y^* be nonnegative vectors representing an underlying ratio scale. They are related to X and Y by

$$\begin{aligned} X &= X^* + C_n; \\ Y &= Y^* + C_m; \end{aligned} \quad (A.1)$$

where C_n is a vector obtained by multiplying an

$n \times 1$ vector of ones by the constant c , and similarly for C_m . The only restriction on c is that

$$c \leq \min(x_1, \dots, x_n, y_1, \dots, y_m), \quad (A.2)$$

since otherwise X^* or Y^* would have negative elements. $I(X)$ is defined in (12) and c^* is defined in (14). The assertion to be proved is that

$$(a) \ I(X) > I(Y) \text{ and } c^* > \min(x_1, \dots, x_n, y_1, \dots, y_m) \geq 0;$$

$$(b) \ I(X) > I(Y) \text{ and } \mu_x I(X) = \mu_y I(Y) \text{ implies}$$

$$(c) \ I(X^*) > I(Y^*).$$

Proof by contradiction. Assume that (a) or (b) is true, and assume for the moment that (c) is false, i.e., $I(X^*) \leq I(Y^*)$. From (A.1) we have

$$I(X - C_n) \leq I(Y - C_m)$$

which, by (13), is

$$\left(\frac{\mu_x}{\mu_x - c} \right) I(X) \leq \left(\frac{\mu_y}{\mu_y - c} \right) I(Y).$$

This reduces to

$$c[\mu_x I(X) - \mu_y I(Y)] \geq \mu_x \mu_y [I(X) - I(Y)]. \quad (A.3)$$

If (b) is true, (A.3) simplifies to $I(Y) > I(X)$ which is a contradiction. Assume (a) is true. The assumption that $c^* > 0$ implies that $\mu_x I(X) - \mu_y I(Y) > 0$. Formula (A.3) then becomes

$$c \geq \frac{\mu_x \mu_y [I(X) - I(Y)]}{\mu_x I(X) - \mu_y I(Y)} = c^* > \min(x_1, \dots, x_n, y_1, \dots, y_m),$$

which contradicts (A.2). Therefore, $I(X^*) > I(Y^*)$.

REFERENCES

- Aigner, Dennis J. and Arthur S. Goldberger
1970 "Estimation of Pareto's law from grouped observations." *Journal of the American Statistical Association* 65: 712-23.
- Aigner, Dennis J. and A. J. Heins
1967 "A social welfare view of the measurement of income equality." *Review of Income and Wealth* 13:12-5.
- Aitchison, J. and J. A. C. Brown
1957 *The Lognormal Distribution*. Cambridge, Eng.: Cambridge University Press.
- Alker, H. and B. M. Russett
1964 "On measuring inequality." *Behavioral Science* 9:207-18.
- Atkinson, Anthony B.
1970 "On the measurement of inequality." *Journal of Economic Theory* 2:244-63.
- Bartels, C. P. A. and P. Nijkamp
1976 "An empirical welfare approach to regional income distributions." *Socio-Economic Planning Sciences* 10:117-28.

- Blau, Peter M.
 1977a *Inequality and Heterogeneity*. New York: Free Press.
 1977b "A macrosociological theory of social structure." *American Journal of Sociology* 83:26-54.
- Champernowne, David
 1974 "A comparison of measures of inequality of income distribution." *The Economic Journal* 84:787-816.
- Cox, D. R. and D. V. Hinkley
 1974 *Theoretical Statistics*. London: Chapman and Hall.
- Creedy, John
 1977 "The principle of transfers and the variance of the logarithms." *Oxford Bulletin of Economics and Statistics* 39:152-8.
- Dalton, Hugh
 1920 "The measurement of the inequality of incomes." *Economic Journal* 30:348-61.
- Dasgupta, Partha, Amartya Sen and David Strett
 1973 "Notes on the measurement of inequality." *Journal of Economic Theory* 6: 180-7.
- Easterlin, R. A.
 1974 "Does economic growth improve the human lot? some empirical evidence." Pp. 89-125 in P. A. David and M. W. Reder (eds.), *Nations and Households in Economic Growth*. New York: Academic Press.
- Farley, Reynolds
 1977 "Trends in racial inequalities." *American Sociological Review* 42:189-208.
- Gastwirth, Joseph L.
 1972 "The estimation of the Lorenz curve and the Gini index." *The Review of Economics and Statistics* 54:306-16.
 1975 "The estimation of a family of measures of income inequality." *Journal of Econometrics* 3:61-70.
- Hall, M. and N. Tideman
 1967 "Measures of concentration." *Journal of the American Statistical Association* 62:162-8.
- Hauser, R. M., P. J. Dickinson, H. P. Travis, and J. M. Koffel
 1975 "Structural changes in occupational mobility among men in the United States." *American Sociological Review* 40:585-98.
- Hays, William
 1963 *Statistics*. New York: Holt, Rinehart and Winston.
- Hewitt, Christopher
 1977 "The effect of political democracy and social democracy on inequality in industrial societies." *American Sociological Review* 42:450-64.
- Jackman, Robert W.
 1974 "Political democracy and social equality: a comparative analysis." *American Sociological Review* 39:29-45.
- Jencks, Christopher, M. Smith, H. Acland, M. J. Bane, D. Cohen, H. Gintis, B. Heyns, and S. Michelson
 1972 *Inequality*. New York: Harper and Row.
- Kakwani, N. C. and N. Podder
 1973 "On the estimation of Lorenz curves from grouped observations." *International Economic Review* 14:278-92.
 1976 "Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations." *Econometrics* 44:137-8.
- Kelley, Jonathan and Herbert S. Klein
 1977 "Revolution and the rebirth of inequality: a theory of stratification in post-revolutionary society." *American Journal of Sociology* 83:78-99.
- Kendall, Maurice and Alan Stuart
 1977 *The Advanced Theory of Statistics*. Vol. 1. 4th ed. New York: Macmillan.
- Kolm, Serge-Christophe
 1976 "Unequal inequalities, I." *Journal of Economic Theory* 12:416-42.
- Lenski, Gerhard
 1966 *Power and Privilege*. New York: McGraw-Hill.
- Martin, J. David and Louis N. Gray
 1971 "Measurement of relative variation." *American Sociological Review* 36:496-502.
- Mehran, Farhad
 1975 "Bounds on the Gini index based on observed points of the Lorenz curve." *Journal of the American Statistical Association* 70:64-6.
- Morris, Carl
 1972 *Measures of Relative Income Inequality*. Santa Monica: Rand.
- Mueller, John H., Karl F. Schuessler and Herbert L. Costner
 1977 *Statistical Reasoning in Sociology*. 3rd ed. Boston: Houghton-Mifflin.
- Nerlove, Marc and S. James Press
 1973 *Univariate and Multivariate Log-Linear and Logistic Models*. Santa Monica: Rand.
- Pyatt, Graham
 1976 "On the interpretation and disaggregation of Gini coefficients." *Economic Journal* 86:243-55.
- Ray, James Lee and J. David Singer
 1973 "Measuring the concentration of power in the international system." *Sociological Methods and Research* 1:403-37.
- Rothschild, Michael and Joseph E. Stiglitz
 1973 "Some further results on the measurement of inequality." *Journal of Economic Theory* 6:188-204.
- Rubinson, Richard and Dan Quinlan
 1977 "Democracy and social inequality: a re-analysis." *American Sociological Review* 42:611-23.
- Salem, A. B. Z. and T. D. Mount
 1974 "A convenient descriptive model of in-

- come distribution: the gamma density." *Econometrica* 42:1115-27.
- Schutz, R. R.
1951 "On the measurement of income inequality." *American Economic Review* 41:107-22.
- Sen, Amartya
1973 *On Economic Inequality*. New York: Norton.
- Theil, Henri
1967 *Economics and Information Theory*. Chicago: Rand McNally.
- Williamson, Jeffrey G.
1977 "Strategic wage goods, prices, and inequality." *American Economic Review* 66:29-41.
- Yntema, D. B.
1933 "Measures of inequality in the personal distribution of income or wealth." *Journal of the American Statistical Association* 28:423-33.

THE EFFECT OF DIRECT GOVERNMENT INVOLVEMENT IN THE ECONOMY ON THE DEGREE OF INCOME INEQUALITY: A CROSS-NATIONAL STUDY*

STEVEN STACK

Alma College

American Sociological Review 1978, Vol. 43 (December):880-888

Sociological theories of income inequality have neglected a Keynesian approach to the problem. This paper tests the Keynesian notion that the degree of direct government involvement in the economy should reduce income inequality through such means as full employment and the facilitation of economic growth. A regression analysis of data from 32 nations indicates that the degree of direct government involvement in the economy is the single most important factor associated with low income inequality. This relationship is independent of both the level of economic development and the rate of economic growth.

The sociological explanation of income inequality has focused on a number of recurrent themes such as economic development theory (Kuznets, 1963; Kerr et al., 1964; Cutright, 1967; Peters, 1973; Hewitt, 1977; Robinson and Quinlan, 1977), and world-economy theory (Cutright, 1967; Robinson, 1976). However, there is also a fourth possible theory, one built upon the Keynesian theory of political economy.¹ The degree of income in-

equality also depends upon the extent to which the state is directly involved in the national economy. Such mechanisms as employment producing expenditures, state provided employment, and the enactment of monetary and fiscal policies that will augment the rate of growth have considerable bearing on the degree of inequality in society. With data from 32 nations, this paper performs a test of two related models of income inequality, the

* Address all communications to: Steven Stack; Department of Sociology; Alma College; Alma, MI 48801.

The author would like to thank Kenneth Neubeck, William V. D'Antonio, Bill Harrell, Gerald Sazama, and the anonymous reviewers for their helpful comments, suggestions, criticisms, and encouragement on earlier versions of this work.

¹ The Keynesian model of income inequality might be viewed as a refinement and extension of the political model. For example, the political model focuses on elements of political organization and development which are presumed to have an effect on income inequality, while the Keynesian model focuses on the actual policies adopted by governments in their efforts to reduce inequality. However, there are

a number of differences between the two perspectives which may serve to separate the two models. Political models of income redistribution tend to view the problem from a conflict perspective wherein the common people obtain universal suffrage and, through such mechanisms as the formation of social-democratic parties, use the state apparatus to obtain a larger share of the income at the expense of the elite. However, in a Keynesian perspective we may understand the process of redistribution through the use of a consensus frame of reference. Here the balancing of savings, consumption and investment benefits the propertied class as well as the common people. Relative to a state of recession, the achievement of such an economic equilibrium increases profits as well as wages.

economic development model and what we will term the *neo-Keynesian model*.

THE MODEL

The first portion of the model of income inequality tested in this paper is based on aspects of economic development theory. The economic development or modernization theory emphasizes the notion that *the logic of industrialization*, and not political factors, determines the nature of the stratification system (Kerr et al., 1964). The principal reason for this standardizing influence of the economic factor has to do with the relationship between industrialization and occupational structure. Increased technological complexity and economic growth create a need for highly skilled manpower. These processes continually increase the number and proportion of high-status and highly skilled positions. At the same time, through automation and other factors, the proportion of low status and low skilled positions declines. There emerges a vast middle in the stratification system. Members of this large class have the necessary bargaining power based on status and extensive training to command sizable wages and salaries. The elite in society must share an increasing proportion of the income with the middle class in order to maintain the supply and flow of skilled workers to man the increasingly complex technology. The gap between the rich and poor diminishes and the degree of income inequality will decline. Kravis (1960:413-5) constructs a similar argument. The increasing differentiation of the occupational structure is at the heart of what might be termed the process of *destratification*.

The rate of economic growth, assumed to be strongly associated with the level of development, provides the second major causal variable in modernization theory. In the early stages of economic development it is generally assumed that economic growth increases inequality, given such conditions as a relatively low level of occupational differentiation, a small class of skilled craftsmen, and a relatively non-democratic political regime. However, in the long run, industrial machine produc-

tion increases the level of the societal surplus so much that the elite can afford to redistribute income to the common people. Machines add such a considerable amount to society's wealth that a new condition arises. The elite can increase the absolute size of its income, while at the same time decreasing its relative share (Lenski, 1966; Cutright, 1967); given that economic growth is maintained, we can expect industrial society to foster reduced inequality.

There has been a number of empirical studies on the relationship between the level of economic development and the degree of income inequality, but the nature of the relationship is inconclusive (Kuznets, 1963; Jackman, 1974; Robinson, 1976; Hewitt, 1977). On the one hand, for example, Hewitt (1977) finds that the level of development, measured by per capita energy consumption, is significantly related to the share in income of the top quintile, even when we control for other variables such as simple democracy, rate of economic growth and socialist party strength. On the other hand, for example, Robinson (1976) finds that the level of development, as measured by the logarithmic transformation of per capita energy consumption, is not significantly related to the share of the top quintile once we introduce controls for variables of the world-economy model. Different measures of each of the two key variables, different samples, and different sets of control variables are among the reasons for the inconsistent findings in this area. It is beyond the scope of this paper to resolve these inconsistent studies. Our purpose is to use the level of economic development and the rate of economic growth as control variables.²

² A preliminary analysis involving 88 nations of the relationship between level of industrialization and the Gini index of inequality found that the relationship resembles a sine curve of the form $Y = B_0 + B_1X_1 + B_2X_1^2 + B_3X_1^3$ (Stack, 1976b). Inequality first increases with advances in technology in pre-industrial society, decreases in early and midindustrialization, but at high levels of industrialization it shows signs of increasing once again. The suggested explanation is that the rate of economic growth, one factor at the heart of why industrialization should reduce inequality, is a full 50% less, on the average, for the most advanced nations as compared with

Keynes's (1964) theory of employment, interest, and money can be applied to the specific problem of income inequality. Keynes's theory is relevant through its concern with the relationship between direct government involvement in the economy and the categories of employment and economic growth. Keynes's theory offers an explanation for the variation in employment and economic growth rates; these, in turn, can be applied to the problem of income inequality (the greater the employment and growth rate, the less the inequality).

Direct government in the economy is viewed in Keynesian theory as improving the probability of achieving the fundamental goal of an equilibrium between saving, consumption, and investment (Leijonhufvud, 1968; Lekachman, 1964). The level of employment, important in determining the level of income inequality, is a function of the demand for goods and services. Demand is, in turn, a function of the relative propensity to consume and the propensity to save. If the amount of money saved by income recipients is greater than the amount required by those who are responsible for investment, then total demand will be insufficient to sustain full employment.³ If too much is saved, there is not enough demand for commodities to stimulate more investment in plants and equipment and the job creation rate slows. In

addition there may be layoffs. Too much saving then is "bad" for the economy in the sense that it will lead to recessions and the problem of unemployment.⁴ We advance the notion that the relative proportion of poor and near poor increases as unemployment rates increase. This results in an increase in the bipolarization of income classes and an increase in the degree of income inequality.

The central government can introduce policies designed to balance savings, consumption and investment. Such *Keynesian* policies include government spending and the manipulation of the money supply. The present paper restricts itself to the former policy.⁵ Revenues derived from such sources as progressive taxes on corporate profits and individual incomes can be used to augment demand through social security programs, welfare expenditures, public works projects, and so on. Such government spending decreases the probability that there will be excessive saving, that there will be unemployment, and that there will be a slow rate of job creation since investors have market incentives to invest. Government spending is a key mechanism for solving the problem of unemployment. In this capacity the central government also can reduce the degree of inequality of incomes.

nations in the midstages of industrialization. However, these results are quite tenuous since the degree of the comparability of the data used by the author has yet to be determined by their source (Shail, 1974).

³ In Keynesian theory too little saving is also a problem for economic organization since the associated problem of excessive wages causes inflationary pressures. In this case some writers argue that high inflation will increase income inequality (Galbraith, 1958:221, 264-5; Kohler, 1968:239). For example, relatively powerless, unorganized groups such as state workers, persons living on social security, pensioners, and welfare recipients, will tend to fall behind the inflationary spiral, whereas organized workers will keep ahead of it, thus expanding income inequality. While it is beyond the scope of the present paper to test this application of Keynesian theory, prior research has found no significant relationship between inflation and income inequality (Stack, 1976a:143). While inflation undoubtedly hurts the economic position of some groups more than others, the evidence suggests that the overall degree of inequality stays the same.

⁴ According to the position of classical economics and "Say's Law," if the propensity to save is excessive, there is a built-in, relatively natural tendency for the economy to tend back towards a state of equilibrium through the fall of interest rates. However, Keynes disagrees with this premise. He argues that the tendency, if any, for the interest rate to fall in a situation of excessive savings depends on the policy of the central bank which is, in turn, somewhat independent of the market forces discussed in classical theory. Even if the central bank voluntarily decreases the interest rate in such a situation, it may do so too late or the decrease may not be enough to ward off a recession.

⁵ It is beyond the scope of a single paper to discuss all of the theoretically relevant Keynesian policies with respect to the problem of income stratification. However, the other major state tool that can be used to regulate the level of unemployment, and in our perspective the associated degree of income inequality, is monetary policy. For example, if we wish the interest rates to be driven down in a situation of excessive savings, a key governmental policy might be to increase the money supply. This may be a key factor in fostering increased investment, greater demand, and hence greater employment opportunities.

Obviously, various forms of government spending, such as defense and highway expenditures, will have differential effects on employment and aggregate demand. For example, as Baran and Sweezy (1966) point out, given the capital intensive nature of military production, the same amount of money spent on labor intensive industries would result in more employment, and we would add, a higher probability of economic equilibrium. In like manner, various forms of government taxation will have differential effects on the after tax distribution of income, but we will not pursue this complex issue in the present investigation.⁶ We anticipate that differences in government spending will be robust enough to influence income inequality despite the variation among nations in the nature of these expenditures.

That the degree of direct government involvement in the economy in the form of government spending is related to the level of unemployment is the subject of recent research (Stack, 1977; 1978c). Direct government involvement was found to be the major factor related to low levels of unemployment both at zero-order level ($r = -.81$) and in the regression analysis ($t = -7.173$, $\beta = -.939$, $p < .05$). In contrast, indicators of the degree of institutionalized political development did not affect unemployment independent of government spending.⁷ The magnitude of these differences for individual nations can be appreciated with reference to the United States. The United States, the na-

tion with the highest rate of unemployment, 6.7%, had relatively low government expenditures totaling 21% of its GNP. In contrast, Sweden, with one of the highest levels of government expenditures, 39.1% of its GNP, had an unemployment rate of only 1.5%. The nearly 100% difference between these two nations in government spending is associated with an over 300% difference in their unemployment rates.

The impact of government spending on unemployment and income inequality should be felt not only through direct job creation programs in government, but also through the *multiplier effect* or the multiple reinvestment of money obtained from higher levels of consumption in job creative ventures in the private sector.

The impact of such job creation through such means as public works projects, government ownership of industry, and the multiplier effect, also augments the rate of economic growth. This, in turn, fosters a climate, as discussed earlier, favoring redistribution since the affluent can reduce their relative share of the income while at the same time increase the absolute amount of real income.

In summation, we anticipate that the greater the direct government involvement in the economy (DGI) the less the unemployment and the less the income inequality. Also, the greater the DGI, the greater the economic growth, and the less the degree of income inequality.

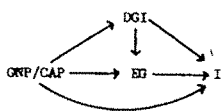
The necessity for government spending is considered to be directly related to the level of economic development (Galbraith, 1967; Schumpeter, 1962). We anticipate, however, that government spending will have an effect on income inequality independent of the level of development.

Our model is summarized in Figure 1. The level of economic development influences income inequality directly through occupational differentiation and indirectly through its direct association with the pace of economic growth. The level of economic development is directly associated with the degree of direct government involvement (DGI). Direct government involvement influences income inequality directly through such mechanisms as providing public works employment and

⁶ A recent investigation of the personal income tax, one of the chief mechanisms of raising money for government spending, found significant variation in the extent to which different nations incorporate a truly effective progressiveness in such taxes which significantly reduces inequality in the after tax distribution of income. The extent to which personal income taxes are truly progressive is almost entirely a political consideration wherein factors such as voter turnout, age of democratic institutions, and the occurrence of acts of organized political violence explain over 80% of the variance in the degree of progressiveness (Stack, 1978b).

⁷ An index of collective protest, or noninstitutionalized political participation, was significantly related to the unemployment rate. To the extent that the degree of income inequality is related to unemployment, this suggests that indicators of noninstitutional politics may be important to future investigations of income inequality.

indirectly through its fostering of economic growth.⁸



Key:

- GNP/CAP = Gross National Product per capita
 DGI = Direct Government Involvement in the Economy
 EG = Rate of Economic Growth
 I = Degree of Income Inequality

Figure 1. The Model

METHODOLOGY

The Dependent Variable

Cutright (1967) and Jackman (1974;1975) use a rather unsatisfactory index of income inequality based on the distribution of income to eight selected economic sectors such as agriculture and manufacturing. This measure was viewed by both writers as a substitute for a more valid and reliable index based on the household personal distribution of income. However, data on the latter measure were not available at the time.

With the work of Adelman and Morris (1971) comparable data on income inequality based on the personal or household distribution of income were made available. Paukert (1973) increased the size of the Adelman and Morris data bank to 56 nations. Paukert omits data for four of their nations due to inadequacies, and, using more realistic assumptions, he recalculates the inequality index of nine na-

tions. He replaces their data on three countries with data from superior sources, and adds comparable data for 16 advanced nations. The present study utilizes Paukert's data bank.⁹

We are concerned with measuring the overall degree of income inequality. A variety of indices has been developed to summarize the degree of total inequality. Most of these are empirically equivalent (Alker and Russett, 1956:358-72). The present study will use the well-known Gini index of inequality.¹⁰

The level of economic development is customarily measured in terms of either GNP/capita or energy consumption/capita (Jackman, 1974; 1975; Hewitt, 1977; Robinson, 1977; Robinson and Quinlan, 1977). The choice of one's basic measure, as well as the choice between the use of a transformation of the data, such as logarithmic transformations, and the use of the raw data, is conditioned largely by the degree to which each measure is related to inequality (Hewitt, 1977:460). Using this criterion, we chose GNP/capita as our measure. As in Hewitt's (1977) research we found that the nontransformed variable provided a better fit for the curve than did the logarithmic version. In a similar fashion, we chose GNP/capita as our measure of the rate of economic growth. Our specific measure is the average rate of growth during the years 1950-1965. The data are taken from Taylor and Hudson (1972:Table 5).

Direct government involvement in the

⁸ Other writers such as Robinson (1976) and Adelman and Morris (1973) also have investigated the role of direct government involvement in the economy on income inequality, but these treatments utilize different theoretical frameworks. Robinson uses a measure of government expenditures as a measure of state strength in world economy involving tariff policies and other means of power, but tariff policies seem a bit far removed from government revenue/GDP, Robinson's index (Stack, 1978a:271). Adelman and Morris also do not connect direct government involvement in the economy to Keynesian theory. Their discussion of this variable, which is just one of 48 in their exploratory study, is largely restricted to the problem of its measurement as opposed to its theoretical relevance (Adelman and Morris, 1973:63-5).

⁹ In addition to the data in the Paukert data bank we were able to obtain comparable data on eight socialist nations from the meticulous work of Michal (1973), Scafnicki (1971), the United Nations (1967), and Yanowitch (1969). Other relatively noncomparable data and/or indices of inequality from scattered sources have always found that socialist nations tend to have less inequality than capitalist ones (Lydall, 1968; Parkin, 1971; Wiles and Markowski, 1971; Pryor, 1971; Schnitzer, 1974; Wiles, 1974; 1975; Chenery et al., 1974; Stack, 1976a; 1976c). Our new data are consistent with these earlier findings; socialist nations have less inequality than the capitalist ones. We would anticipate this from the standpoint of our model since socialist command economies are well-known for their high level of direct government involvement in the economy (Carson, 1973).

¹⁰ For a detailed description of the nature and method of calculation of the Gini coefficient see Bronfenbrenner (1971:45-53).

economy is measured in terms of the level of government spending. Comparative data on government spending were collected as part of the cross-national data archive of Russett et al. (1964). We will utilize their data on the sum of the expenditures of the central government, social security, and public enterprises as a percent of the GNP. These data are for the year 1960. The alternative would be to use the revenue of the governmental institutions. This would be less desirable since it omits deficit spending from the category of expenditures. Our index includes social security expenditures which augment demand for goods and services, the expenditures of publically-owned enterprises which provide direct employment opportunities, and general government expenditures which prime the economy. The sum of these expenditures is standardized for the size of the national economy by dividing by the gross national product for the year 1960. Data were available on this measure for 32 nations for which income inequality data were also available. Our study is based on these countries. Our findings must be taken with great caution due to the small number of nations under investigation and because the set of nations under investigation is not a sample in any technical sense.

THE ANALYSIS

It has been hypothesized that direct government involvement in the economy, the level of development, and the rate of economic growth all should be negatively related to the degree of income inequality. To test these and our associated hypotheses on the relationships between our independent variables, we first computed a matrix of zero-order Pearson product-moment correlation coefficients which is given in Table 1. Three of the five hypothesized relationships, as diagrammed in Figure 1, meet the requirements for statistical significance at at least the .05 level. The variable with the strongest zero-order correlation with income inequality is direct government involvement ($r = -.74$, $p < .005$). In addition, as predicted, the rate of economic growth is negatively related to income inequality ($r = -.46$, $p < .005$).

Table 1. The Matrix of Zero-Order Pearson Product-Moment Correlation Coefficients

Variable	GINI	DGI	GNP/CAP	EG
GINI ¹	1.000	-.7437**	-.2448	-.4585**
DGI ²		1.000	.0102	.6092**
GNP/CAP ³			1.000	-.0578
EG ⁴				1.000

¹ Gini: Gini index of income inequality circa 1960.

² DGI: degree of direct government involvement in the economy as measured by the expenditures of the central government, social security, and public enterprises as a percentage of the GNP, 1957.

³ GNP/CAP: GNP per capita, 1965.

⁴ EG: rate of economic growth, 1950-1965.

** Statistically significant at the .005 level.

The Keynesian notion that direct government involvement increases the rate of economic growth also is supported ($r = .69$, $p < .005$). The negative relationship between the level of economic development and inequality is in the expected direction but just misses statistical significance at the .05 level. This may be due to its slight negative association with economic growth which may be suppressing the negative relationship. The positive association between economic growth and direct government involvement may render one of these variables' negative associations with income inequality spuri-

Table 2. Effects of Direct Government Involvement, GNP/Capita, and the Rate of Economic Growth on the Degree of Income Inequality (N=32)

Variable	Regression Coefficient	Standard Error of Coefficient	Computed Value of Student's t	Beta Coefficient
DGI	-0.003	0.001	-4.839**	-0.721
GNP/CAP	-0.000	0.000	-2.021*	-0.239
EG	-0.002	0.010	-0.221**	-0.333
Intercept	.543168			
R ² =	.57			
F=	14.60, 3/28 df			

See notes to Table 1.

* Statistically significant at the .05 level.

** Statistically significant at the .005 level.

ous in the regression analysis. That is, for example, if we control for DGI, the relationship between inequality and economic growth may disappear, indicating growth influences inequality only through its association with direct governmental involvement. At this rudimentary level of analysis we can tentatively conclude that direct government involvement in the economy is the factor most closely associated with low income inequality.

In order to weigh the relative importance of the independent variables and to test for spurious and suppressed relationships a multiple regression analysis was performed. The results of the regression analysis are given in Table 2. The *t*-statistics associated with DGI and GNP/capita indicate that the relationships with the dependent variable are in the expected direction and are statistically significant. If we control for economic growth and level of development, there is a negative and significant relationship between DGI and income inequality ($p < .001$). Also, if we control for DGI and economic growth, the greater the level of development the less the inequality. However, if we control for DGI and level of development, there is no significant relationship between the rate of economic growth and inequality. This indicates that the original association between growth and inequality is spurious and that growth affects inequality indirectly through its association with DGI.

To weigh the importance of the two significant independent variables, we can take the absolute values of their beta coefficients (Kleinbaum and Kupper, 1978: 170-1). The greater the absolute value of a variable's beta coefficient, the greater its importance in explaining the variation in the dependent variable. In our analysis, the degree of direct government involvement in the economy is the most important variable associated with inequality of incomes ($\text{beta} = -.721$). Per capita GNP is the second most important variable ($\text{beta} = -.239$). We also calculated the elasticity for DGI. A 1% increase in DGI is associated with a .36% decrease in income inequality.¹¹

¹¹ The concept of elasticity has been used quite frequently in business and economics. It is defined as

Taken together, the variables explained 61% of the variance in inequality. However, the R^2 should be adjusted for the problem of shrinkage due to our small sample size (Ferguson, 1976:465-6). We recalculated our R^2 to adjust it for this problem.¹² The adjusted R^2 indicates that the equation explains 57% of the variance. The significance of the equation also can be demonstrated by the *F* ratio of 14.6 (3,28 df, $p < .01$).

CONCLUSION

This paper performed a test of a Keynesian model of income inequality. The results indicate support for the model that are independent of the level of industrialization. Caution should be exercised, however, in interpreting these results due to the small number of nations under investigation. The greater the direct government involvement in the economy the less the income inequality. We suggest that the link between DGI and low income inequality is due to the former's facilitation of full employment, not to its zero-order association with economic growth which was determined to be spurious by the regression analysis.

A final question is how well does the neo-Keynesian model of income inequality compare with its major alternatives, the political model and the world-economy model? Great caution must be

the percentage change in the dependent variable given a percentage change in the independent variable. The elasticity of a variable is calculated as: $E =$

$$B_j \frac{\bar{X}_j}{\bar{Y}}$$

where: \bar{X} is the mean of the independent variable;
 \bar{Y} is the mean of the dependent variable;
 B_j is the regression coefficient of X_j .

Some caution should be exercised in interpreting elasticity coefficients since their value will depend upon the point used. An investigator may choose some point other than the variable means for theoretical or policy reasons. This is especially true if there is a wide range in the data (Hantshek and Jackson, 1977:79; Pindyck and Rubinfeld, 1976:72-3).

¹² The formula for adjusting R^2 for the problem of shrinkage is taken from Ferguson (1976:465) and is:

$$\bar{R}^2 = 1 - \frac{N - 1}{N - g - 1} (1 - R^2).$$

exercised in drawing an answer to this question due to such problems as the possibility of relationships between major causal variables and differences in the samples of nations in the relevant major studies. However, we tentatively propose that the neo-Keynesian model is more powerful than the world-economy model with respect to the amount of variance explained. Robinson (1976) does not report the amount of variance explained by his model, but since *none* of his zero-order correlations between factors of world economy and overall income inequality are greater than + or -.331, and since there is a problem of multicollinearity between several of his factors of world economy, we anticipate that the amount of variance explained is considerably less than in our model. Hewitt (1977:462) reports that his political model explains 58% of the variance in the share of income of the top 5% and top 20% of the population. This compares favorably with our neo-Keynesian model.¹³ Further research might explore a synthesis of the two models.

APPENDIX

Nations in the Analysis

Argentina	Greece	Poland
Australia	Hungary	Romania
Bolivia	India	South Africa
Brazil	Iraq	Sweden
Burma	Israel	United Kingdom
Bulgaria	Italy	United States
Czechoslovakia	Japan	U.S.S.R.
Denmark	Netherlands	Venezuela
East Germany	Norway	West Germany
Finland	Peru	Yugoslavia
France	Philippines	

REFERENCES

- Adelman, Irma and Cynthia Taft Morris
1971 *An Anatomy of Patterns of Income Distribution in Developing Nations*. Part 3 of the Final Report (Grant AID/csd-2236).
- 1973 *Economic Growth and Social Equity in Developing Countries*. Stanford: Stanford University Press.
- Alker, Hayward and Bruce M. Russett
1966 "Indices for comparing inequality." Pp. 349-72 in Richard L. Merritt and Stein Rokkan (eds.), *Comparing Nations: The Use of Quantitative Data in Cross-National Research*. New Haven: Yale University Press.
- Baran, Paul and Paul M. Sweezy
1966 *Monopoly Capital*. New York: Monthly Review Press.
- Blalock, Hubert M., Jr.
1972 *Social Statistics*. New York: McGraw-Hill.
- Bronfenbrenner, Martin
1971 *Income Distribution Theory*. Chicago: Aldine.
- Carson, Richard
1973 *Comparative Economic Systems*. New York: MacMillan.
- Chenery, Hollis, Montek Ahluwalia, C. L. G. Bell, John H. DuLoy, and Richard Jolly
1974 *Redistribution with Growth*. London: Oxford University Press.
- Cutright, Phillips
1967 "Inequality: a cross-national analysis." *American Sociological Review* 32:562-78.
- Edwards, Allen L.
1976 *An Introduction to Linear Regression and Correlation*. San Francisco: Freeman.
- Ferguson, George
1976 *Statistical Analysis in Psychology and Education*. New York: McGraw-Hill.
- Galbraith, John K.
1958 *The Affluent Society*. Boston: Houghton Mifflin.
1967 *The New Industrial State*. New York: Signet.
- Hanushek, Eric and John Jackson
1977 *Statistical Methods for the Social Sciences*. New York: Academic Press.
- Hewitt, Christopher
1977 "The effect of political democracy and social democracy on equality in industrial societies: a cross-national comparison." *American Sociological Review* 42:450-64.
- Jackman, Robert
1974 "Political democracy and social equality." *American Sociological Review* 39:29-45.
1975 *Politics and Social Equality*. New York: Wiley.
- Kerr, Clark, J. T. Dunlap, Frederick Harbison and Charles Myers
1964 *Industrialism and Industrial Man*. New York: Oxford University Press.
- Keynes, John
1964 *The General Theory*. New York: Harcourt Brace.
- Kleinbaum, David and Lawrence Kupper
1978 *Applied Regression Analysis and Other Multivariate Methods*. North Scituate: Duxbury Press.
- ¹³ However, there is some reason to believe that the neo-Keynesian model is somewhat superior to the political model on statistical grounds. Since Hewitt's (1977) study is based on a smaller sample ($N = 18$) and since he has more predictor variables in his regression analysis, we would anticipate a larger R^2 (Blalock, 1972:468). When the number of variables begins to approach the number of cases in a regression analysis, we can expect to obtain very large multiple correlations simply because we can take advantage of chance fluctuation. Hewitt (1977) violates one statistical rule of thumb by having more than one variable for every ten cases in his sample (Edwards, 1976; Hanushek and Jackson, 1977).

- Kohler, Heinz
1968 *Scarcity Challenged*. New York: Holt, Rhinehart, and Winston.
- Kravis, Irving
1960 "International differences in the distribution of income." *Review of Economics and Statistics* 42:408-16.
- Kuznets, Simon
1963 "Quantitative aspects of the economic growth of nations, 8: the distribution of income by size." *Economic Development and Cultural Change* 11:1-80.
- Leijonhufvud, Axel
1968 *On Keynesian Economics and the Economics of Keynes*. New York: Oxford University Press.
- Lekachman, Robert
1964 *Keynes' General Theory*. New York: St. Martin's.
- Lenski, Gerhard
1966 *Power and Privilege*. New York: McGraw-Hill.
- Lydall, Harold
1968 *The Structure of Earnings*. New York: Oxford University Press.
- Michal, Jan
1973 "Income distribution of earnings and household incomes in small socialist nations." *Review of Income and Wealth* 19:407-28.
- Parkin, Frank
1973 *Class Inequality and the Political Order*. New York: Praeger.
- Paukert, Felix
1973 "Income distribution at different levels of development: a survey of evidence." *International Labor Review* 108:97-125.
- Peters, B. Guy
1973 "Income inequality in Sweden and the United Kingdom: a longitudinal analysis." *Acta Sociologica* 16:108-20.
- Pindyck, Robert S. and Daniel Rubinfeld
1976 *Econometric Models and Economic Forecasts*. New York: McGraw-Hill.
- Pryor, Frederic L.
1971 "Economic system and the size distribution of income and wealth." *International Development Research Center, Indiana University, Bloomington*.
- Rubinson, Richard
1976 "The world-economy and the distribution of income within states: a cross-national study." *American Sociological Review* 41:638-59.
- Rubinson, Richard and Daniel Quinlan
1977 "Democracy and social inequality: a reanalysis." *American Sociological Review* 42:611-23.
- Russett, Bruce M., Hayward R. Alker, Jr., Karl W. Deutsch, and Harold D. Lasswell
1964 *World Handbook of Political and Social Indicators*. New Haven: Yale University Press.
- Scafnicki, K.
1971 "Distribution of wages and income among individuals and groups." *Polish Sociological Bulletin* 24:32-47.
- Schnitzer, Martin
1974 *Income Distribution: A Comparative Study of the United States, Sweden, West Germany, East Germany, the United Kingdom, and Japan*. New York: Praeger.
- Schumpeter, Joseph A.
1962 *Capitalism, Socialism, and Democracy*. New York: Harper and Row.
- Shail, Jain
1974 *Size Distribution of Income: Compilation of Data*. Washington, D.C.: International Bank for Reconstruction and Development.
- Stack, Steven
1976a *Inequality in Industrial Society: Income Distribution in Capitalist and Socialist Nations*. Ph.D. dissertation, Department of Sociology, University of Connecticut.
1976b "Income stratification and the level of socio-economic development: an empirical test of Lenski's theory." Paper presented at the annual meeting of the Eastern Sociological Society, Boston.
1976c "Income stratification and the type of economic system: a cross-national test of 44 countries." Paper presented at the annual meeting of the American Sociological Association, New York.
1977 "Multivariate analysis of the political economy of unemployment." *Peninsular Papers: Journal of the Michigan Sociological Association* 2:18-33.
1978a "Internal political organization and the world-economy of income inequality (comment on Rubinson, ASR August, 1976)." *American Sociological Review* 43:271-2.
1978b "Political organization and the degree of tax progressiveness." *International Modern Review of Sociology* 8. In press.
1978c "Unemployment: a comparative analysis." *Proceedings of the Eighth Annual Alpha Kappa Delta Sociological Research Symposium, Richmond*. In press.
- Taylor, Charles and Michael C. Hudson
1972 *World Handbook of Political and Social Indicators*. New Haven: Yale University Press.
- United Nations
1967 *Economic Survey of Europe in 1965, Part 2; Incomes in Postwar Europe; A Study of Policies, Growth, and Distribution*. Geneva: United Nations Press.
- Wiles, Peter
1974 *Distribution of Incomes: East and West*. Amsterdam: North Holland.
1975 "Recent data on Soviet income distribution." *Survey* 21:28-41.
- Wiles, Peter and Stefan Markowski
1971 "Income distribution under communism and capitalism: some facts about Poland, the UK, the USA, and the USSR." *Soviet Studies* 22:344-70; 487-511.
- Yanowitch, Murray
1969 "The Soviet income revolution." Pp. 143-54 in Celia Heller (ed.), *Structured Social Inequality*. New York: Macmillan.

PRODUCTIVITY AND ACADEMIC POSITION IN THE SCIENTIFIC CAREER*

J. SCOTT LONG

Washington State University

American Sociological Review 1978, Vol. 43 (December):889-908

This paper examines the interrelationship between scientific productivity and academic position, two key dimensions of the scientific career. Contrary to the results of most earlier studies, the effect of departmental location on productivity is found to be strong, whereas the effect of productivity on the allocation of positions is found to be weak. Productivity, as indicated by measures of publications and citations, is shown to have an insignificant effect on both the prestige of a scientist's initial academic appointment and on the outcome of institution changes later in the career. Although the relationship between productivity and the prestige of an academic appointment is insignificant at the time a position is obtained, the effect of departmental prestige on productivity increases steadily with time. For those scientists who change institutions, the prestige of the new department significantly affects changes in a scientist's productivity after the move. It is argued that past studies have obtained spurious results due to their failure to employ a longitudinal design. Not only do cross-sectional designs provide misleading results regarding the interrelationship between departmental location and productivity, but they also systematically alter the findings regarding the effects of sponsorship and doctoral training on productivity.

Two fundamental questions underlie past studies of the academic careers of scientists. First, what accounts for the unequal distribution of prestige in the academic stratification system? Why do some scientists obtain positions in highly prestigious graduate departments while others find themselves in relatively unknown departments? Second, what accounts for the great inequality in the productivity of scientists? If the institutional goal of science is the extension of certified knowledge (Merton, 1942:267-85), why do some scientists succeed, publishing widely-acclaimed work year after year, while other scientists publish almost nothing throughout their careers? Natural ability, dedication, quality of education, and sponsorship have all been considered as explanations of this unequal distribution of scientific productivity and aca-

ademic position. In addition, location in the stratification system and productivity have each been considered as dependent upon the other.

While productivity has been examined as a primary causal influence on the allocation of positions in the scientific community, and departmental location has been considered as a factor explaining the inequality in productivity, the interdependence of these two factors over the course of the scientific career has not been systematically explored. This is an important limitation since if these factors do have a reciprocal relationship, then an empirical analysis of either factor which fails to control for the other must yield biased results. This bias will not only affect conclusions about the nature of the relationship between productivity and academic position, but will also alter findings on the effects of other variables upon productivity and position.

In this paper the relationship between productivity and position is examined by studying the early careers of Ph.D. biochemists employed in U.S. graduate departments. By employing a longitudinal design, the effect of productivity on the allocation of position can be isolated from

* Address all communications to J. Scott Long; Department of Sociology; Washington State University; Pullman, WA 99164.

I wish to thank R. McGinnis, P. Allison, E. Hackett, L. Hargens, and B. Reskin for their comments on an earlier version of this paper. Support was provided by NIH Contract NO1-OD-4-2175 and NIMH Training Grant 5T32MH14244, awarded to Cornell's Research Program on Social Analysis of Science Systems, Robert McGinnis, Director.

the effect of departmental location on productivity. We find that while the effect of productivity on the prestige of a position is weak, the effect of departmental location on productivity is strong. These findings are contrary to much recent empirical work in this area; we suggest that the failure of past studies to utilize a longitudinal design accounts for this disagreement.

PRIOR RESULTS

Prior studies have consistently reported a moderate positive relationship between scientific productivity and departmental prestige (Cole and Cole, 1973; Crane, 1965; Hagstrom, 1968; Hargens and Hagstrom, 1967; Manis, 1951). This is true even though these studies differed in research design, examined a variety of fields, and employed different measures of scientific productivity (based both on publications and citations to publications) and prestige (based on both departmental and institutional location).

Two distinct explanations of the correlation between position and productivity have been offered. First, it has been argued that more prestigious departments *select* more productive scientists for their faculty; productivity is causally prior to position. Alternatively, it has been argued that being at a more prestigious department *facilitates* greater productivity; position is causally prior to productivity. In the remainder of this section we will review these alternative explanations.

Productivity and Departmental Selection

If prestigious departments select more productive scientists for faculty positions, a correlation between a scientist's productivity and the prestige of his academic position should be found. Empirical work along this line has attempted to isolate those factors which determine the prestige of a scientist's academic appointment. Historically such research falls into two groups: the first consisting of studies of the academic profession pursued in large part out of curiosity, concern or practical interest in how the academic marketplace

is functioning; the second motivated by the work of Robert Merton on the normative structure of science, where the recruitment process is examined as an instance of the allocation of rewards by the scientific community. Researchers in the first group have consistently seen little effect of productivity on position, while researchers in the second group have, to varying extents, found a positive effect.

The best example of the first set of studies is Caplow and McGee's (1958) *The Academic Marketplace*. After examining the process by which a number of academic vacancies were filled, they concluded that hiring decisions were not based on actual evaluation of the applicant's work, but rather were based on the prestige of the candidate's graduate department, the eminence of his sponsors, and chance. Logan Wilson (1942:29, 40, 50) similarly noted the importance of the prestige that the young graduate obtains from his affiliation with his doctoral institution. And Max Weber (1919:132) warns students that chance looms large in the careers of scientists; achievement does not necessarily lead to deserved position. Thus, studies of the academic profession consistently emphasize the importance of the scientist's perceived prestige, prestige based more on the established reputations of his sponsors and doctoral department than on his actual scientific work.

The second group of studies, dealing with the social functioning of the science system rather than the characteristics of the academic profession, are founded on Merton's (1973) seminal work on the normative structure of the social system of science. It is argued that science is governed by the norm of universalism, according to which careers should be open to talent and not be determined by functionally irrelevant characteristics. Since the production of certified knowledge is the institutionally sanctioned goal of science, rewards should be allocated on the basis of an individual's contribution to the body of scientific knowledge. Appointment to a prestigious position in the hierarchy of academic departments, which is considered to be a form of recognition (Cole and Cole, 1973:57), should be

allocated to reward a scientist's productivity.

Several studies have examined this hypothesis. Cole and Cole (1973:98) found that a scientist's productivity has as strong an effect on academic appointment as does the ranking of the scientist's doctoral department. Similarly, Hargens and Hagstrom (1967) found that the effects of doctoral origins and productivity on current academic affiliation are approximately equal for their entire sample; the effect of doctoral origins is stronger for younger scientists, however. Crane (1970), on the other hand, concluded that the prestige of the doctorate has more influence than productivity upon selection for a prestigious position. She suggested (1970:961) that "despite the system's normative commitment to universalistic criteria, they are not utilized in practice."

Departmental Effects on Productivity

Less frequently the relationship between productivity and position has been considered as the result of the effects of departmental location on scientific productivity. A variety of mechanisms for a departmental effect have been proposed (cf. Crane, 1970; Hagstrom, 1968): prestigious departments provide a scientist with more free time for research, superior physical resources, abler research assistants, more stimulating colleagues, and stronger social support for doing research; and there may be a "halo" effect for being located at a prestigious department whereby papers and grant proposals appear more impressive to reviewers and to potential citers.

Crane (1965) concluded that the effects of departmental position on productivity are positive, but smaller than those of doctoral origin. Unfortunately, in her study no attempt was made to determine if the relative strengths of these effects change (with departmental effects becoming stronger) over the course of the scientist's career. Hagstrom (1967:61) concluded that while departmental prestige is associated with a number of factors which would be expected to influence productivity, there is no support for "believing that the greater productivity of scientists in high prestige departments is due more

to the context in which they work, the facilities and assistants that are available, than to their research skills and motivations."

Overall, then, no consensus emerges from past studies. Little evidence is found for a departmental effect, although tests of such an effect are relatively rare. And while the support for the importance of productivity in the selection process is stronger, substantial conflicting evidence also exists.

METHODOLOGICAL NOTE

In evaluating these studies it is important to realize that they are all limited by their designs—none is longitudinal, monitoring the work and publication histories of scientists over an extended period of time. As a result, the causal ordering of productivity and position is inherently ambiguous. Studies which explain position in terms of productivity would have been equally justified on empirical grounds to turn their analysis around and to have explained productivity in terms of position. A longitudinal research design is necessary to distinguish between the competing hypotheses of a departmental effect vs. a selection effect. In the absence of such a design, estimates of the effect of productivity on position will be upwardly biased to the extent that a departmental effect is operating, a point made by Hargens and Hagstrom (1967:30).

The following analysis employs a longitudinal design in order to untangle the interrelationships between productivity and position. The selection process is examined in two ways. First, the movement of a scientist into his first academic job from either graduate school or a postdoctoral fellowship is analyzed as a function of characteristics of his education and his productivity up to the time of his appointment. Second, the movement of a scientist from one institution to another is studied. The relative strengths of the effects of educational experiences, productivity in the period immediately prior to the move, and the prestige of the first position are assessed.

The effect of departmental location on productivity also is tested in two ways.

First the change in the productivity of a scientist after he obtains his first position is related to the prestige of his location. By examining the changes in productivity levels of those scientists who do *not* change positions, the possibility of a selection process causing the change in correlations between productivity and position can be eliminated. Second, changes in the productivity of those scientists who do change institutions are analyzed. If departmental location is a factor affecting an individual's scientific productivity, a change in productivity should be observed after departmental mobility and this change should be related to the prestige of the new department. We conclude by summarizing our findings and relating them to previous research.

DATA AND MEASUREMENT

Analysis is based on the population of male biochemists who obtained their doctorates in the fiscal years 1957, 1958, 1962 and 1963.¹ From this population of Ph.D.s two groups are isolated. The first and larger group consists of 134 biochemists who held teaching positions in graduate departments evaluated by the Roose and Andersen (1970) study and who were classified as *nonmovers*. To be considered a nonmover, a biochemist had to have held a position in the same department from the time he entered the academic marketplace (which may have occurred after a postdoctoral fellowship) until our time series ends (from ten to fifteen years after the doctorate, depending upon the year of its receipt). In order to insure adequate time series for uncovering departmental effects, the set of nonmovers also was limited to those for whom six consecutive years of data were available. The second group (N=47) consists of biochemists who held teaching positions in two different departments in two different institutions during the first ten to fifteen years of their careers. To be included

in this set of *movers*, a biochemist had to have held his position in the first location for at least three years and his position in the second location for at least four years before our time series ended. Those scientists who met these conditions but who left their teaching positions for nonteaching activities (e.g., university administration, research in industry or government) were excluded since earlier analysis (Long, 1977a:124-6) suggests that the productivity patterns of such scientists prior to their moves out of academia are atypical.

Our measure of departmental prestige is based on the Roose and Andersen (1970) study.² Since biochemists obtain teaching positions in a variety of departments other than biochemistry (e.g., chemistry, physiology) and accurate information on the actual department in which an individual was located is often impossible to obtain, an index of the overall prestige of those departments in which a biochemist might be teaching was constructed. This bioscience prestige index was defined as the weighted sum of the prestige scores of an institution's departments of biochemistry (1/12), chemistry (1/4), physiology (1/12), microbiology/bacteriology (1/12), and pharmacology (1/12), where the weights (contained in parentheses) were based on estimates of the number of biochemists teaching in each field. These scores range from 100 to 500.

Productivity was measured using counts of both publications and citations to them. *Chemical Abstracts* (1955-1973) was used to locate the articles by the biochemists in our sample. Citations to these articles, as found in the *Science Citation Index* (Institute for Scientific Information, Vols. 1961, 1964, 1966, 1968, 1970, 1972, and 1974), were used for constructing the citation measures. The name of the first author on multiple-authored papers where the cohort member was not the first author was used to locate citations to junior-authored papers; thus downward bias in counts for scientists who were predominantly junior authors was avoided (see Reskin, 1977:494-5, for

¹ Analysis was restricted to male biochemists due to the small number of females who obtained doctorates in biochemistry during this period, and to the difficulty in obtaining complete information on those females who did obtain degrees.

² These data were kindly provided by Charles J. Andersen.

details of this problem). From these publication and citation counts, measures were constructed for each year the biochemists were located in rated departments. For a given year, the publication measures reflect three years of publications and the citation measures are restricted to citations to publications in this three-year period. While in most studies citation measures cover all past publications this modification was introduced to make the measure of citations more sensitive to changes in the utilization of an individual's work. Since coverage by *Science Citation Index* and *Chemical Abstracts* increased during the period covered by our study, counts were standardized within years of the Ph.D. After standardization the scale of the measures was adjusted back to the scale of the raw counts and the minimum value was reestablished as zero. For purposes of analysis, the measures were transformed by the square root to compress the upper end of the original metric. Exact details on the construction of yearly citation measures are presented in the appendix.

Finally, a series of measures to reflect the biochemists' educational experiences were coded. Since approximately two-thirds of the biochemists held postdoctoral fellowships,³ the prestige of the postdoctoral institution was coded using the bioscience index described above. A measure of the prestige of the last educational location was constructed by assigning a biochemist the prestige score of his doctoral department if he did not have a postdoctoral fellowship in a graduate department, and the prestige of the fellowship institution if he did. This was done since earlier analysis (Long, 1977b) indicates that for postdoctoral fellows the prestige of the postdoctoral location replaces the prestige of the doctoral department in affecting career outcomes.

A measure of the mentor's accom-

plishments proximate to the period when the sample members obtained their degrees was constructed by counting citations to the mentor's articles published in the five years prior to 1961. This measure will be referred to variously as the "mentor's eminence," the "mentor's prestige," or the "recognition received by the mentor." It should be kept in mind, however, that citations are not an unambiguous indicator, since they may reflect both the performance of a scientist and his or her standing or prestige in the scientific community. Distinguishing between these two dimensions of a mentor's possible influences on the careers of his students is an important topic, but one that is beyond the scope of this paper (cf. Reskin, 1978).

Collaboration with the mentor was indicated by a dummy variable assigned a value of one if the mentor coauthored on one or more articles published before or during the year after the student's receipt of the doctorate, and zero otherwise. The year after the Ph.D. was included in the determination of predoctoral collaboration to reflect the lag between the completion of research and its publication.

Astin's (1971) score for the selectivity of an institution was assigned to each baccalaureate institution. The selectivity index has values ranging from one to seven, with seven being the most selective category. This variable, which has been interpreted as a crude indicator of intelligence and as a measure of the quality of baccalaureate education, has been shown to be a moderately successful predictor of future success.

PRODUCTIVITY AND THE ALLOCATION OF ACADEMIC POSITIONS

To assess the magnitude of the effect of productivity on appointment to a prestigious position, productivity must be measured at the time a scientist obtains that position, not several years later. For if productivity measures are based on work completed in the department of the appointment, the effect of departmental location on productivity may bias the estimates. This is demonstrated in Table 1, which presents the results of regressions on the prestige of the initial academic ap-

³ While two-thirds of the biochemists studied in this paper began their careers with postdoctoral fellowships, it should be understood that among the entire cohort (i.e., both those with academic and nonacademic careers) the fraction having immediate postdoctoral fellowships is one-half. This fraction is in agreement with estimates from other sources (cf. Curtis, 1969).

Table 1. Regressions Relating Predoctoral Statuses and Productivity to Prestige of First Academic Position, for Those 134 Biochemists Who Do Not Change Institutions

Equation		Coefficients* of					R ²	d.f.
		ORIG	MENT	SEL	PUB	CIT		
1. Prestige								
Regressed on	β	.370	.202	.145	-.101	.152	.304	127
Productivity at	t	4.66	2.48	1.89	1.01	1.49		
Job Year 1	r	.464	.372	.249	.096	.204		
2. Prestige								
Regressed on	β	.295	.173	.128	.015	.270	.358	127
Productivity at	t	3.72	2.23	1.75	0.16	2.78		
Job Year 7	r	.464	.372	.249	.327	.446		

NOTE: Dependent variable is the Roose-Andersen bioscience prestige score of the institutional location. Item identifications are: ORIG=origin prestige, Cartter prestige of Ph.D. department if non-fellow or if nonacademic fellow and Roose-Andersen bioscience prestige score if fellow; MENT=square root of five year citations counts to mentor; SEL=selectivity of baccalaureate institution; PUB=publication level, square roots of standardized levels of three-year publication counts ending in the given year; CIT=citation level, square roots of standardized levels of citations to publications in the three-year period ending in the given year.

* Row β gives standardized regression coefficients; row t gives the t-statistics (with more than 120 degrees of freedom in the regression, critical values for a two-tailed test of significance at the .10, .05 and .01 levels are: 1.645, 1.960 and 2.576, respectively; for a one-tailed test the critical values are: 1.282, 1.645 and 2.326 for significance levels of .10, .05 and .01, respectively); row r gives zero-order correlation coefficients with dependent variable.

pointments for the 134 biochemists described above as nonmovers.⁴

Equation 1 regresses departmental prestige on characteristics of the scientist's training and productivity. The productivity measures are based only on work completed before the scientist obtained his first position.⁵ Thus the correlations between productivity and position, .096 for publications and .204 for citations, cannot be the result of a departmental effect. These correlations are substantially smaller than those reported in the studies reviewed above.⁶ Further, the standardized regression coefficients from publications and citations are small and statistically insignificant, while those from the prestige of the sponsoring department

(either the doctoral or fellowship department), the mentor's eminence, and the selectivity of the baccalaureate institution are strong, positive and statistically significant. Thus no support is found for the importance of productivity in the recruitment process.⁷

The importance of using productivity measures based only on work completed at the time of the appointment can be seen in equation 2 of Table 1. This equation includes productivity measures based on work completed in the three-year period ending in the sixth year of the scientist's location at the department whose prestige is the dependent variable. The standardized regression coefficient from citations to the prestige of the academic appointment has increased substantially from .152 to .270. The coefficient is now

⁴ The process by which biochemists obtain their initial academic positions has been examined in detail elsewhere (Long, 1977b). Here we only report those results which have the greatest impact on our argument concerning the relative importance of selection and departmental effects.

⁵ To take into account the time lag between the completion of research and its publication in a professional journal, we included research published during the scientist's first year at his new position in these productivity measures.

⁶ Crane (1970) is an exception which will be considered below.

⁷ Scientists who held their first position in the department where they obtained their doctorate or pursued postdoctoral studies (i.e., inbred scientists) have been included in these and later analyses. When these scientists are excluded, the coefficient from the origin prestige is attenuated but the other coefficients remain essentially the same. Substantive conclusions are not affected by the inclusion of inbred scientists in the reported results (cf. Hagstrom and Hargens, 1968).

the largest in the equation and it has attained statistical significance. Further, the magnitudes of the coefficients from the prestige of the doctoral origin and the citation level are now approximately equal to those reported by Cole and Cole (1973:98) for a similar regression.⁸ But clearly equation 2 is not the appropriate test of the selection hypothesis since it "explains" departmental prestige on the basis of work completed at that department, and thus confounds the selection effect with the departmental effect. Yet this is essentially what is being done in cross-sectional studies that relate the prestige of a scientist's location at a given time to his productivity at that time, not at the time he obtained the position.

Note also that the apparent increase in the effect of productivity is associated with an attenuation of the effects of other variables, even though the zero-order correlations for these variables are necessarily unchanged. Thus the regression effect of the prestige of the sponsoring department decreases from .370 to .295, the effect of the mentor's prestige decreases from .202 to .173, and the effect of the baccalaureate institution decreases from .145 to .128. So, not only is the productivity effect overestimated in a cross-sectional design, but the effects of other variables are underestimated. In short, changing from a cross-sectional to a longitudinal design results in substantively and theoretically different conclusions.

Before rejecting the operation of a selection effect, a further complication must be considered. It is plausible that departments hiring young scientists do not base their selection on the demonstrated scientific contributions of the candidates since the young scientists have not had adequate time to establish their scientific

merit. Professionally older scientists, who have had the time to publish a body of research, may be evaluated and recruited on the basis of their published work. Hargens and Hagstrom (1967:34) suggest that in this way the stratification system reduces the initial inequalities in the allocation of position.

Our data do not support this argument. First, in an analysis not reported here (see Long, 1977b) it is found that for biochemists who begin their careers with postdoctoral fellowships the delayed process of job allocation does not differ significantly from the allocation process for those obtaining positions immediately after receiving the doctorate. Since the fellows have devoted several additional years to research prior to entering the job market,⁹ one would expect on the basis of the above hypothesis that their recruitment would be based more closely on their productivity.

Second, the institution changes which occur later in the career do not seem to be affected by productivity. Table 2 presents regressions on the outcomes of 47 institution changes occurring on average 7.6 years after the doctorate.¹⁰ In equation 1 the destination prestige of the move is regressed on the prestige of the scientist's initial academic department, characteristics of his training, and his productivity at the time of the move. The effect of citations on the prestige of the academic appointment has increased slightly from equation 1 of Table 1, and with more observations we would expect the coefficient to become statistically significant. Still, the magnitude of the effect remains small and does not offer strong support for the hypothesis of the universalistic allocation of positions later in the career. Indeed, the scientist's productivity at the

⁸ It is generally argued that only unstandardized regression coefficients should be compared across populations. But, as Hargens (1976) has argued, in cases where it is theoretically reasonable to expect standardized coefficients to be structural parameters, the comparison of standardized coefficients is justified. Certainly in the case of evaluating a scientist's productivity it is not his number of publications or citations that is important, but rather his number relative to other scientists in his field. Accordingly, the comparison of standardized coefficients is appropriate in this case.

⁹ On average the postdoctoral fellows in our sample began their first tenure-track faculty position 3.2 years after the doctorate, with 60% of the fellows having positions lasting three or more years.

¹⁰ While the results on institution changes are limited by the small number of cases, we note that similar findings were obtained in analyses of 72 institution changes (which include these 47 cases). The results presented here are limited to 47 cases since only these cases provide sufficient time series for testing the impact of departmental location on productivity (analysis presented below).

Table 2. Regressions Relating Predoctoral Statuses, Productivity and Prestige of First Institutional Location to Prestige of Second Institutional Location for Those 47 Biochemists Who Held a Position at the First Institution for at Least Three Years and at the Second Institution for at Least Four Years

Equation	Coefficients* of							R ²	d.f.
		PRST	ORIG	MENT	SEL	PUB	CIT		
1. Prestige									
Regressed on									
Productivity at	β	-.173	.081	.121	.105	-.111	.195	.069	40
the Time of the	t	0.98	0.48	0.75	0.66	0.52	0.94		
Institution Change	r	-.088	.093	.124	.102	-.001	.102		
2. Prestige									
Regressed on									
Productivity Three	β	-.175	.002	.076	.027	-.012	.387	.167	40
Years after the	t	1.04	0.00	0.49	0.18	0.05	1.86		
Institution Change	r	-.088	.094	.124	.102	.195	.365		

NOTE: Dependent variable is the Roose-Andersen bioscience prestige score of the second institutional location. Item identifications are: PRST=prestige of first institutional location, Roose-Andersen bioscience prestige score; ORIG=origin prestige, Cartter prestige of Ph.D. department if nonfellow or if nonacademic fellow and Roose-Andersen bioscience prestige score if fellow; MENT=square root of five-year citation counts for mentor; SEL=selectivity of baccalaureate institution; PUB=publication level, square roots of standardized levels of three-year publication counts ending in the given year; CIT=citation level, square roots of standardized values of citations to publications in the three-year period ending in the given year.

* Row β gives standardized regression coefficients; row t gives the t-statistics (with 40 degrees of freedom in the regression, critical values for a two-tailed test of significance at the .10, .05 and .01 levels are: 1.684, 2.021 and 2.704 respectively; for a one-tailed test the critical values are: 1.303, 1.684 and 2.423 for significance levels of .10, .05 and .01, respectively); row r gives the zero-order correlations with the dependent variable.

time of the move is more strongly related to the prestige of the origin institution (zero-order correlations of .457 with publications and .382 with citations) than with the prestige of the destination institution (zero-order correlations of -.001 with publications and .102 with citations). Later career mobility does not operate in such a way as to create greater correspondence between the prestige of a scientist's academic affiliation and his productivity.

Somewhat surprisingly, no other variables in equation 1 of Table 2 have a significant relationship to the prestige of the destination institution. The prestige of the doctoral origin no longer aids the scientist in obtaining a position. The mentor appears to continue to affect his student's placement, but the positive effect is greatly attenuated. Finally, there is no strong relationship between the prestige of the first institution of employment and the prestige of the destination institution.

There is some evidence that the tradeoff

between academic rank and institutional prestige, and the negative effects of age influence the outcomes of institutional mobility. Caplow and McGee (1958:148) suggest that downward mobility is often accompanied by advancement in rank. For all biochemists who change institutions during the period our data covers (which includes the 47 cases in Table 2), 62% of the scientists moving to a less prestigious department advanced in academic rank, while only 37% of those moving to a more prestigious department advanced in rank. Caplow and McGee (1958:44, 186-7) also suggest that older scientists are professionally less attractive since their future productivity may be less than that of younger scientists. And indeed, we find that the older the scientist the less likely he is to move to a more prestigious institution.

Still, none of the factors affecting the outcome of institution changes has a strong effect. This suggests that the mechanism by which job changes occur is

different from that determining the initial deployment of scientists, where doctoral origins and characteristics of the mentor have a significant effect. This inability to uncover factors that strongly influence job changes may be related to Caplow and McGee's (1958:47, 53, 55, 77, 82, 91, 160) finding that a wide variety of both personal and professional factors influence a scientist's decision to change institutions. The uniqueness of these factors cannot be tapped with the variables in our study.

These findings may be compared with two similar analyses. Crane (1970) examined the recruitment of faculty members in six departments among the 20 most prestigious graduate schools. While she did not report regression coefficients predicting the destination prestige and combined those who obtained their first position with those who changed institutions, she indicated that the correlations with destination prestige and the number of publications is $-.051$ and the correlation with the number of citations is $-.001$, with these measures being based on work completed *before* the move. She also reported that productivity had no effect on the prestige of the position obtained.

Allison (1976:205-18) analyzed the institution changes of 894 scientists from Hagstrom's (1967) study of four fields. Unlike Crane's and our findings, productivity emerged as a significant factor affecting destination prestige. But, as Allison (1976:185) warned, his productivity measure is based on citations made *after* the job changes. Hence if changes in productivity which occur after the job change are systematically related to the prestige of the destination department, his coefficient for the effect of productivity will be confounded with the effect of departmental location on productivity. Results presented in equation 2 of Table 2 indicate how this may occur. In this equation productivity measures are based on work completed after the job change. Both the correlations of productivity with destination prestige and the coefficient from citations increase substantially. The utilization of the scientist's work (as indicated by citations to it) appears to influence significantly the outcome of his institution change. Thus, given Allison's warning

about his productivity measures and the findings of equation 2, we can explain the difference between his results and ours.

Overall there is no evidence of the importance of productivity in the recruitment of faculty. Rather, in the initial placement of graduates the prestige of the sponsoring department and the eminence of the mentor play major roles. Institution changes occurring later in the career appear to be unaffected by either characteristics of the education of the scientist or by his productivity. Moreover, if productivity measures which are based on work completed after appointment to a department are used to predict the prestige of that department, it will *appear* as though productivity is a causally significant factor. These findings suggest the importance of a departmental effect, an issue which is now considered.

EFFECTS OF DEPARTMENTAL LOCATION ON PRODUCTIVITY

Tables 3 through 6 present evidence for the existence of a departmental effect on productivity. Tables 3 and 4 examine the effects of departmental location on publications and citations to publications¹¹ for those 134 biochemists who do *not* change institutions during the period for which we have data. Accordingly, changes in the relationship between departmental prestige and productivity cannot be attributed to the awarding of more prestigious positions to those who are productive. Tables 5 and 6 examine the effects of departmental location on those who change institutions during the period for which we have data. This allows us to explore the effects of changes in departmental location on productivity.

For scientists who do not change departments, publication levels are not immediately affected by departmental prestige, as is seen in equations 1 and 2 of Table 3. Equation 1 shows that produc-

¹¹ Recall that the publication measure is based only on publications in a three-year period ending in the given year. Similarly, the citation measure is based only on citations to publications in this three-year period, not to all past publications. This was done to make the productivity measure more sensitive to changes over time.

Table 3. Regressions Relating Predoctoral Statuses, Institutional Location and Earlier Productivity to Publication Level, for Those 134 Biochemists Who Do Not Change Institutions

Equation		Coefficients* of						R ²	d.f.
		PRST	ORIG	MENT	COL-LAB	SEL	PUB (Yr-3)	CIT (Yr-3)	
1. Publication Level at Job Year 3	β	-.038	-.005	.239	.135	.034			.084 127
	t	0.38	0.05	2.50	1.55	0.39			
	r								
2. Publication Level at Job Year 3	β	-.046	-.026	.184	.047	.005	.211	.153	.179 125
	t	0.48	0.27	1.99	0.54	0.05	1.83	1.32	
	r	.066	.078	.255	.177	.055	.359	.348	
3. Publication Level at Job Year 6	β	.220	.181	-.039	.171	.103			.170 127
	t	2.29	1.93	0.43	2.06	1.22			
	r								
4. Publication Level at Job Year 6	β	.252	.180	-.169	.095	.087	.593	-.049	.463 125
	t	3.14	2.37	2.22	1.40	1.27	6.82	0.54	
	r	.327	.300	.157	.192	.166	.572	.393	

NOTE: Dependent variables are square roots of standardized three-year publication counts ending in the given year. Item identifications are: PRST=prestige of institutional location, Roose-Andersen bioscience prestige score; ORIG=origin prestige, Cartter prestige of Ph.D. department if nonfellow or if nonacademic fellow and Roose-Andersen bioscience prestige score if fellow; MENT=square root of five-year citation counts for mentor; COLLAB=collaboration with mentor, defined as one if any article published by the year after the doctorate was coauthored with the mentor, zero otherwise; SEL=selectivity of baccalaureate institution; PUB(Yr-3)=publication level three years earlier, square roots of standardized levels of three-year publication counts ending three years before the given year; CIT(Yr-3)=citation level three years earlier, square roots of standardized levels of citations to publications in the three-year period ending three years before the given year.

* Row β gives standardized regression coefficients; row t gives the t-statistics; row r gives zero-order correlations with dependent variable.

tivity levels are affected by the prestige of the scientist's mentor during his doctoral study and whether or not he collaborated with his mentor. By controlling for earlier productivity, in equation 2, insights into how the mentor influences his student's later productivity may be obtained. Later in the career collaboration does not directly affect productivity, but rather operates through its effects on the early productivity of the student. For at the time of doctoral studies, collaboration with the mentor accounts for a substantial proportion of the student's publications. Even though actual collaboration with the mentor usually ceases after receipt of the degree, the influence of early publications continues to affect the student's rate of publication. The strong positive effect of early publications on later publications indicates a tendency of those who publish to continue publishing. There remains, however, a good deal of variation over time in individual levels of productivity as indicated by the correlation of only .359 between publication levels in consecutive

three-year periods. While the effect on publications of the mentor's prestige remains significant, it is somewhat attenuated by the introduction of controls for early productivity. This, as in the case of collaboration, results from a positive, albeit weaker, effect of the mentor's prestige on the student's predoctoral productivity.

Equations 3 and 4 show that departmental prestige significantly influences the scientist's productivity after the third year in a department. In equation 3 we find that departmental location is the strongest factor explaining the scientist's level of publications. Note that this effect persists after controlling for earlier productivity, as seen in equation 4. Thus *change* in the publication rate is most strongly affected by the prestige of the scientist's academic location—scientists in prestigious departments increase their publication rate while those in less prestigious departments begin to publish relatively less.

As departmental effects emerge, the effects of variables associated with the

biochemist's doctoral education change. The effects of the prestige of the doctoral origin on the level of publications and on the change in the publication rate increase substantially. The total effect operates in two ways. First, the prestige of the doctorate acts indirectly through its effect on the prestige of the scientist's academic appointment, which in turn affects the scientist's publication rate. Second, the prestige of the doctorate directly affects the biochemist's later publication pattern. This direct effect would be consistent with the argument that students at better institutions are more talented or better trained, and that in time this talent or training emerges through greater productivity. An alternative hypothesis, however, is that the greater number of informal contacts which students in prestigious departments have the opportunity to make eventually pays off in resources that facilitate productivity. While the effect of collaboration with the mentor on the level of publications increases, this remains an

indirect effect operating through the effect of early productivity on later productivity, as was the case in equations 1 and 2. Finally, the effect of the mentor's eminence on changes in the student's rate of publication later in the career has become *negative*, of moderate strength, and statistically significant. This is an intriguing and unanticipated finding which indicates that after a scientist's productivity pattern has been established (in this case, six years into the teaching career), those scientists with more prestigious mentors begin to publish less. The fact that the effect of the mentor's eminence on the citations received by a scientist is insignificant (Table 4, equation 4), suggests that scientists with more eminent mentors begin to be more selective in their publishing practices, publishing less frequently, *ceteris paribus*, than other scientists, but maintaining their total number of citations.

Results for citation levels are presented in Table 4. As with publications, the departmental effect is strong and positive.

Table 4. Regressions Relating Predoctoral Statuses, Institutional Location and Earlier Productivity to Citation Level, for Those 134 Biochemists Who Do Not Change Institutions

Equation	Coefficients* of								R ²	d.f.
		PRST	ORIG	MENT	COL-LAB	SEL	PUB (Yr-3)	CIT (Yr-3)		
1. Citation Level at Job Year 3	β	.192	-.068	.239	.074	.097			.147	127
	t	1.97	0.71	2.59	0.89	1.14				
2. Citation Level at Job Year 3	β	.178	-.080	.164	-.003	.040	.056	.343	.273	125
	t	1.96	0.90	1.87	0.03	0.49	0.51	3.14		
	r	.279	.125	.317	.124	.170	.332	.455		
3. Citation Level at Job Year 6	β	.335	.123	-.063	.122	.087			.243	127
	t	3.65	1.37	0.73	1.54	1.09				
4. Citation Level at Job Year 6	β	.280	.147	-.080	.061	.045	.264	.338	.501	125
	t	3.62	2.00	1.10	0.94	0.69	3.15	3.89		
	r	.445	.324	.269	.167	.190	.500	.591		

NOTE: Dependent variables are square roots of standardized levels of citations to publications in the three-year period ending in the given year. Item identifications are: PRST=prestige of institutional location, Roose-Andersen bioscience prestige score; ORIG=origin prestige, Cartter prestige of Ph.D. department if nonfellow or if nonacademic fellow and Roose-Andersen bioscience prestige score if fellow; MENT=square root of five-year citation counts for mentor; COLLAB=collaboration with mentor, defined as one if any article published by the year after the doctorate was co-authored with the mentor, zero otherwise; SEL=selectivity of baccalaureate institution; PUB(Yr-3)=publication level three years earlier, square roots of standardized levels of three year publication counts ending three years before the given year; CIT(Yr-3)=citation level three years earlier, square roots of standardized levels of citations to publications in the three-year period ending three years before the given year.

* Row β gives standardized regression coefficients; row t gives t-statistics; row r gives zero-order correlations with dependent variable.

The effect of the scientist's departmental prestige on the utilization of his work (as indicated by citations to it), begins more rapidly after the doctorate and is of greater magnitude than was the case with publications. The citations received by work completed in the first three years at the department are significantly affected by the prestige of the department (equation 1); this effect is maintained after controlling for earlier productivity (equation 2). The departmental effect increases so that by the sixth year at the department the effect on the level of citations (equation 3) is .335, the strongest coefficient in the regression, and the effect of departmental prestige on the change in citations received (equation 4) has increased to .280 and is only slightly less than the effect of early citations on later citations. We may conclude that as a group scientists at prestigious departments produce more highly cited work, regardless of the quality of their education, the prestige of their sponsors, or their earlier productivity.

As with publications, the effects of the mentor are largely indirect after the scientist obtains his doctorate. The correlations between the recognition received by the mentor and the student's citation level decrease with time. By controlling for earlier productivity the effect is shown to be indirect, as was the case with publications. By year six the direct effect has been completely eliminated. Similarly, the effect of collaboration operates indirectly through the impact of collaboration on early publications and citations, and through the effects of this earlier productivity on later productivity. Thus, equations 2 and 4 show that while the correlations with collaboration are in fact increasing, the direct effects remain small. The effect of doctoral prestige on citations increases with time and persists after controlling for earlier productivity, as was the case with publications.

Finally, while there is somewhat more stability in citations than in publications, a great deal of variation over time persists. The correlation between the citations received in the three-year periods ending in years four and seven is only .591. This is less than the figures reported in other

analyses (cf. Cole and Cole, 1973:32). The smaller stability is probably due to our construction of citation measures which reflect the utilization of work *published* in different time periods; thus the effect of the same papers receiving citations in several consecutive years was eliminated. These results are interesting in and of themselves, indicating that a great deal of variability exists in the number of citations a scientist's work receives over time, even though the total number of citations (i.e., citations to all past work) he receives in a given year may be quite stable.

So far our results have shown that controlling for training and earlier productivity, departmental location has a substantial effect on the quantity and utilization (as indicated by citations) of a scientist's work. It is of course possible that the increasing correspondence between productivity levels and the prestige of a scientist's departmental location reflects the hiring institution's ability to judge a scientist's merit by means other than his published work. Thus, if a scientist has only a dissertation and a few publications, or perhaps none at all, at the time he applies for a job, it may be that less formal indicators are evaluated by the hiring department and the departmental effect we uncovered above merely reflects the validation of the department's initial judgments.

To test further our hypothesis of a departmental effect, we examine the productivity of scientists who change institutions.¹² If environment does influence productivity, we would expect the institution changes of scientists to be followed by changes in their productivity which reflect their new environments. Tables 5 and 6 present the results of regressions on productivity levels of 47 scientists who change departments.

Equations 1 and 2 of Table 5 show that while at the first employing (origin) department the effect of departmental location on the scientist's publication level continues to increase until the time of the move. In equation 3 we see that this effect persists after controlling for earlier pro-

¹² This approach was developed in collaboration with Paul Allison. Together we plan a more detailed study of a much larger sample of job changes.

Table 5. Regressions Relating Predoctoral Statuses, Institutional Locations and Prior Productivity to Publication Levels, for Those 47 Biochemists Who Held a Position at the First Institution for at Least Three Years and at the Second Institution for at Least Four Years

Equation	Coefficients* of										R ²	d.f.
	PRST1	PRST2	ORIG	MENT	COLLAB	SEL	PUB (Yr-3)	CIT (Yr-3)				
1. Publication Level Two Years Before Move	β	.124	.423	-.123	.322	.019					.288	41
	t	0.87	2.95	0.86	2.30	0.14						
	r	.265	.414	.088	.276	.105						
2. Publication Level First Year at Second Job	β	.363	.211	-.017	.126	.101					.271	41
	t	2.53	1.46	0.11	0.89	0.73						
3. Publication Level First Year at Second Job	β	.301	.079	.053	.072	.084	.025	.288			.339	39
	t	2.09	0.50	0.36	0.49	0.62	0.11	1.33				
	r	.457	.314	.124	.154	.218	.391	.454				
4. Publication Level Fourth Year at Second Job	β	.362	.069	.095	.233	-.000					.286	40
	t	2.49	0.48	0.65	1.64	0.00						
5. Publication Level Fourth Year at Second Job	β	.103	-.096	.134	.143	-.075	.559	.185			.626	38
	t	0.88	0.85	1.21	1.34	0.71	4.01	1.36				
	r	.392	.206	.243	.288	.162	.719	.572				

NOTE: Dependent variables are square roots of standardized three-year publication counts ending in the given year. Item identifications are PRST1=prestige of first institutional location, Roose-Andersen bioscience prestige score; PRST2=prestige of second institutional location; Roose-Andersen bioscience prestige score; ORIG=origin prestige, Cartier prestige of Ph.D. department if nonfellow or if nonacademic fellow and Roose-Andersen bioscience prestige score if fellow; MENT=square root of five-year citation counts for mentor; COLLAB=collaboration with mentor, defined as one if any article published by the year after the doctorate was coauthored with the mentor, zero otherwise; SEL=selectivity of baccalaureate institution; PUB(Yr-3)=publication level three years earlier, square roots of standardized levels of three-year publication counts ending three years before the given year; CIT(Yr-3)=citation level three years earlier, square roots of standardized levels of citations to publication in the three-year period ending three years before the given year.

* Row β gives standardized regression coefficients; row t gives t-statistics; row r gives zero-order correlations with dependent variable.

ductivity. Thus, while those scientists who change institutions are located at the first institution, the effects of departmental prestige on publications operates in much the same way as it did for those scientists who did not change institutions.

At the time of the move the correlation between publications and the destination prestige is $-.001$, while the correlation with the origin prestige is $.457$. By the fourth year at the new institution, the effect of the old department on the publication level at this time is greater than the effect of the new department when prior productivity is not being controlled (equation 4). But when controls for earlier productivity are introduced (equation 5), the effect of the origin department becomes greatly attenuated and is no longer statistically significant. The effect of the destination prestige, however, decreases only slightly and becomes significant at the $.05$ level. Thus, the *change* in the publication level which occurs after departmental mobility is more closely related to the prestige of the new department than to that of the old department.

Similar findings are presented for citations in Table 6. Comparison of equations 1 and 2 shows that once more the effects of departmental location on citation levels increase before the move. Also, the effect of departmental location is maintained after controlling for earlier productivity. At the time of the move the correlation between the origin prestige and the levels of citations is $.382$, while the correlations with the destination prestige is $.102$. By the fourth year at the new location, however, the effect of the old department's prestige has decreased substantially to $.102$, while the effect of the new department's prestige is $.329$. Further, by comparing equations 4 and 5 we see that the effect of the origin department on later productivity is completely eliminated by controlling for earlier productivity, while the effect of the new departmental environment stays strong and statistically significant.

SUMMARY AND CONCLUSIONS

Our findings may be summarized briefly as follows:

1. Neither the number of publications a biochemist has produced nor the number of citations he has received significantly affects the prestige of his initial academic appointment. Rather, factors related to his graduate education, sponsorship, and postdoctoral study have the greatest influence on his initial academic placement.

2. While there is only a small correlation between productivity and position at the start of a biochemist's employment in an academic department, a stronger relationship emerges through time. This departmental effect occurs independently of the reinforcing effects of earlier productivity on current productivity. Departmental prestige has a much stronger and more immediate effect on the citations a scientist's subsequent work receives than on the quantity of his publications.

3. Later in the career when departmental mobility occurs, the effects of productivity on the prestige of the new employing department increase slightly, but remain quite weak.

4. For scientists who change institutions, the effects of the first department's prestige on productivity increase until the move occurs, at which point they begin to decrease. At the same time, while the correlation between productivity and the new department's prestige is negligible when the position is obtained, the correlations increase markedly within five years. Further, the changes in productivity levels occurring after the move appear to be directly related to the new department's prestige and unrelated to the old department's prestige.

Thus, our longitudinal analysis provides consistent support for the hypothesis that productivity is facilitated by departmental location. Contrary to earlier findings, productivity, as indicated by measures of publication and citation, plays an insignificant role in the selection process. The differences between our results and those of earlier studies may be explained by the failure of past studies to employ longitudinal controls.

Our findings on the placement process

Table 6. Regressions Relating Predoctoral Statuses, Institutional Locations and Prior Productivity to Citation Levels, for Those 47 Biochemists Who Held a Position at the First Institution for at Least Three Years and at the Second Institution for at Least Four Years

Equation	Coefficients* of										R ²	d.f.
		PRST1	PRST2	ORIG	MENT	COLLAB	SEL	PUB (Yr-3)	CIT (Yr-3)			
1. Citation Level Two Years Before Move	β	.204		.422	-.232	.160	.058				.280	41
	t	1.43		2.93	1.61	1.13	0.41					
	r	.330	-.129	.422	-.054	.094	.121					
2. Citation Level First Year at Second Job	β	.282		.270	-.149	.103	.114				.226	41
	t	1.91		1.81	0.99	0.71	0.80					
3. Citation Level First Year at Second Job	β	.256		.197	-.120	.058	.109	.108	.064		.244	39
	t	1.67		1.17	0.77	0.37	0.75	0.46	0.28			
	r	.382	.102	.320	-.006	.085	.191	.324	.340			
4. Citation Level Fourth Year at Second Job	β	.102	.329	.255	.006	.066	.188				.283	40
	t	0.70	2.40	1.75	0.04	0.46	1.34					
5. Citation Level Fourth Year at Second Job	β	-.062	.296	.136	.056	.008	.138	.215	.287		.430	38
	t	0.43	2.34	0.98	0.41	0.06	1.06	1.25	1.70			
	r	.198	.365	.327	.168	.095	.274	.453	.503			

NOTE: Dependent variables are square roots of standardized levels of citations to publications in the three-year period ending in the given year. Item identifications are: PRST1=prestige of first institutional location, Roose-Andersen bioscience prestige score; PRST2=prestige of second institutional location, Roose-Andersen bioscience prestige score; ORIG=origin prestige, Cartier prestige of Ph.D. department if nonfellow or if nonacademic fellow and Roose-Andersen bioscience prestige score if fellow; MENT=square root of five-year citation counts for mentor; COLLAB=collaboration with mentor, defined as one if any article published by the year after the doctorate was coauthored with the mentor, zero otherwise; SEL=selectivity of baccalaureate institution; PUB(Yr-3)=publication level three years earlier, square roots of standardized levels of three-year publication counts ending three years before the given year; CIT(Yr-3)=citation level three years earlier, square roots of standardized levels of citations to publications in the three-year period ending three years before the given year.

* Row β gives standardized regression coefficients; row t gives t-statistics; row r gives zero-order correlations with dependent variable.

are consistent with such studies as those by Caplow and McGee (1958) and Wilson (1942). Studies by Cole and Cole (1973), Hargens and Hagstrom (1967), and Allison (1976) appear to show a positive effect of productivity on positional prestige which is approximately equal to the effect of the doctoral prestige. But this difference may be due to an important assumption these latter studies have made. Hargens and Hagstrom (1967:30) characterize it in this way:

For our purposes, we will adopt the working assumption that a scholar's productivity rate precedes his placement in the academic stratification system and is not greatly affected by his placement. Insofar as a scholar's placement does affect his productivity rate, this assumption will lead us to overestimate the effect of his productivity (and thus of his scholarly merit) in the determination of his location in the stratification system.

All of these studies have made this assumption. Further this would explain the exceedingly low correlation which Crane observed. For in her study she obtained moves occurring in academic years 1963–1965 and used bibliographic sources from the period prior to 1964. Since the majority of Crane's cases involved instances of institutional mobility, as opposed to initial placement, her correlations of prestige with productivity should be approximately equal to those in Table 4, which is in fact the case. Our findings of departmental effects also would be consistent with Hargens and Hagstrom's (1967) findings that the effects of productivity were stronger for the older scientists than the younger. They explained this as the emergence of stronger effects of productivity as a larger body of research became available for evaluation. Alternatively it may reflect the fact that the younger members of their sample had been at their current department for a shorter period of time before productivity was measured, and thus the strength of the departmental effect would be smaller. Similarly, Hagstrom and Hargens (1968) find that the effect of citations on current prestige increases in four path models dealing with steadily older cohorts. While these earlier studies, with the exception of Crane's, do not allow the precise determination of

how long the sample members were at their current departments before productivity was measured, these explanations appear convincing.

The strength of the departmental effect which we find does not contradict Hagstrom's (1967:61) conclusion that there is no basis for "believing that the greater productivity of scientists in high prestige departments is due more to the context in which they work, the facilities and assistants that are available, than to their research skills and motivations." For, as he notes, the fact that one does not find strong intervening factors which explain the importance of prestige upon productivity does not eliminate the possibility that other, unmeasured characteristics are important (Hagstrom, 1968:12). And, as Hagstrom also notes, a longitudinal study showing the emergence of the relationship between productivity and departmental affiliation is necessary. That is what we have provided.

While several important conclusions have been drawn from our analysis, our final understanding of the relationship between productivity and departmental position remains unsatisfactory. Just how does departmental prestige facilitate productivity? Several hypotheses can be constructed, although we do not have sufficient evidence with which to substantiate any of them. First, superior departments may have the prescience to select those who will become (even though they are not now) productive. This would, to some extent, salvage the hypothesis that positions are allocated on the basis of contribution—by adding the qualification, *potential* contribution. Yet it seems unlikely that when departmental mobility occurs (which as Caplow and McGee [1958] have indicated occurs for a myriad of nonprofessional reasons) members of the new department realize that the hired individual's level of productivity will change. Second, the relationship between productivity and departmental location may be due to the effects of departmental characteristics other than prestige. For example, prestigious departments tend to be larger and hence may provide more opportunities for collaboration, which may increase the scientist's overall pro-

ductivity.¹³ Prestige is only one aspect of the context in which a scientist works. It is possible that the departmental effect we have found *underestimates* the global effect of departmental location on productivity. Third, there may be a bias in the evaluation system whereby papers from more prestigious departments *appear* superior (see Zuckerman and Merton, 1971).

Each of these hypotheses will require substantial work to confirm or reject. Given the evidence for a departmental effect which we have presented, inquiry into the above hypotheses is an important next step in studying the scientific career.

Our findings have major significance along two dimensions, one methodological and one substantive. The results presented in Tables 1 and 2 may be considered as examples of what Schoenberg (1977) referred to as *dynamic misspecification*: the application of static models to dynamic processes. Our results have clearly shown that cross-sectional designs can result in biased estimates of structural parameters. In our particular example, the effects of productivity on the allocation of position are inflated, while effects of characteristics of educational experience are attenuated. This substantive example, which illustrates the mathematical conclusions of Schoenberg's work, indicates that great care must be exercised in the interpretation of cross-sectional analysis. Since this is a general problem, not simply relevant to the process of job allocation, it is possible that results of cross-sectional analyses of other dimensions of the exchange process in science may be biased due to dynamic misspecification.

Substantively our findings call for a reconsideration of operation of the reward system in the scientific career. Merton (1973:293) has described the universalistic operation of the exchange process in science as follows:

When the institution of science works efficiently—and like other social institutions, it does not always do so—recognition and esteem accrue to those who have best fulfilled their roles, to those who have made original contributions to the body of scientific knowledge.

Cole and Cole (1973:235) have argued that the major violation of the norm of universalism involves the process of accumulative advantage: "People who have done well at time 1 have a better chance of doing well at time 2, independently of their objective role-performance; the initially successful are given advantage in subsequent competition for rewards." Our results suggest that those who receive prestigious positions at time 1 have a better chance of doing well (in terms of productivity) at time 2, independently of their earlier productivity. But this process differs in one important respect from the process described by Cole and Cole. Advantage accumulates not to those who have been successful at time 1, but to those who have received the advantage of a prestigious position for reasons independent of their productivity at time 1. To the extent that the eminence of a scientist's mentor and the prestige of his doctoral department, *independently* of the productivity of the scientist, are particularistic criteria for evaluation, a *particularistic* advantage accumulates, not an advantage initiated for universalistic reasons. Rather than scientists who have earned their rewards at time 1 having an advantage at time 2, scientists who have had an advantage at time 1 multiply this advantage at time 2. In short, the reinforcing process of accumulative advantage appears to begin with a position obtained in time 1 for particularistic rather than universalistic reasons. Even if success in later job mobility is based more on objective criteria of productivity, the initial academic appointment, which is independent of earlier productivity, has a major impact on later productivity and hence the prestige of the second institutional location. Initial academic appointment significantly affects the chances of a scientist to become a successful researcher. Academic departments may recruit on the basis of the prestige of the mentor and the

¹³ Preliminary analysis has been completed to determine the relationship between collaboration and the departmental effect. This was done by constructing productivity measures in which each publication or citation was divided by the number of authors on the paper. While the departmental effect was slightly attenuated in regressions using these adjusted measures, it remained strong and statistically significant.

Table A1. Algorithm for Creating Yearly Citation Measures from Seven Volumes of *Science Citation Index* *

Year of Citation Measure	SCI-61							SCI-64							SCI-66				
	55	56	57	58	59	60	61	59	60	61	62	63	64		62	63	64	65	66
1955	+							.							.				
1956	+	+						.							.				
1957	+	+	+					.							.				
1958		+	+	+				.							.				
1959			+	+	+			.							.				
1960				+	+	+		.							.				
1961					+	+	+	.	+	+	+				.				
1962						+	+		+	+	+				.				
1963										+	+	+			.				
1964			+	+	+		.	+	+	+	
1965															.		+	+	+
1966															.		+	+	+
1967																			
1968																			
1969																			
1970																			
1971																			
1972																			
1973																			
1974																			

doctoral department because they have insufficient evidence of the young scientist's productivity. But nonetheless, this initial decision to hire, based on where one studied and with whom, has a major effect on the career of the scientist.

APPENDIX

The Construction of Yearly Citation Measures

Longitudinal analysis of the scientific career requires that yearly (or at least periodic) measures of productivity be available. For publication measures this is straightforward: each year of *Chemical Abstracts* (or some appropriate abstracting service) is searched for publications by the scientists under study; alternatively, the source index of *Science Citation Index* (SCI) can be used for the years after 1960. Yearly citation measures pose a greater problem since citation coding is costly and SCI was only first published in 1961. This appendix presents the procedures we have used to deal with these problems in constructing yearly citation measures.

Citations to each article written by members of our sample, whether they were the first author of the article or not, were coded from seven volumes of SCI: 1961, 1964, 1966, 1968, 1970, 1972, and 1974.

From this information citation measures were constructed to reflect the citations received by articles published in a given three-year period. Since the same article can be cited in many years, a decision had to be made regarding which citing years to use for constructing the measures. Table A1 depicts the scheme used. Vertically the year for which a citation count was to be constructed is listed. Horizontally are listed the volumes of SCI which were coded; associated with each volume of SCI is a listing of years in which the majority of articles cited in that volume had been published. In order to make the citation measures sensitive to changes in the rate in which an author's new work is being cited, each yearly count was restricted to publications for a three-year string of articles. A plus (+) in the body of Table A1 indicates that citations to articles from the appropriate year (as indicated by the years at the top or bottom of the table) out of the appropriate volume of SCI were summed to construct the yearly citation measure. In certain cases (e.g., the constructed count for 1964) citations from two volumes of SCI were used. This was done to take into account the half-life phenomenon in citations to articles: articles require time after publication before they can be cited; then, on the average, the number of citations increases before beginning to decline. In biochemistry substantial citation begins one to two years after publication. For example, for the year 1964 a measure of citations to articles published in 1962-1964

Table A1.—Continued

SCI-68					SC-70					SCI-72					SCI-74					Year of Citation Measure
64	65	66	67	68	66	67	68	69	70	68	69	70	71	72	70	71	72	73	74	
																				1955
																				1956
																				1957
																				1958
																				1959
																				1960
																				1961
																				1962
																				1963
																				1964
																				1965
+	+	+																		1966
	+	+	+																	1967
		+	+	+	+	+	+													1968
						+	+	+												1969
							+	+	+	+	+	+								1970
										+	+	+								1971
											+	+	+							1972
												+	+	+	+	+	+			1973
																+	+	+		1974

VOLUME OF SCIENCE CITATION INDEX AND YEAR OF ARTICLE CITED

* Note: Each cell represents citations in a given volume of SCI to articles published in a given year. + 's indicate citations in that cell are used in three-year counts.

was desired. Since SCI volume 1964 was compiled too close to the year 1964 to have many citations to articles from that year, and since SCI volume 1966 was far enough away from articles published in 1962 that citations had begun to fall off, citations from both the 1964 and 1966 volumes of SCI were combined.

Since our sample members received their doctorates in a number of years, it was necessary to standardize these constructed counts by the year of the sample members' Ph.D. Thus, for example, counts from all sample members who received their doctorates in 1957 were standardized to have a mean of zero and a variance of one. There were two major reasons why this standardization was necessary. First the coverage of SCI increases substantially from 1961 to 1974. For example, four years into their careers scientists receiving their doctorates in 1961 would have citation counts from a volume of SCI which surveyed approximately one and one-half times more journals than the corresponding volume of SCI for those receiving their degrees in 1957. It could thus appear that the 1962 cohort produced more cited articles than the 1957 cohort simply because a larger set of articles was surveyed for SCI.

Second, the average number of publications produced by biochemists increases rapidly at the start of

their careers. Due to the half-life of citations, being a year closer to the volume(s) of SCI used to construct the citation counts could make an appreciable difference in the magnitude of the citation measure. To avoid the possibility of giving a particular cohort of scientists an advantage in receiving citations simply because of the temporal relationship of their doctoral year to the years of SCI we used, we standardized the counts by year of the doctorate as described above.

After standardization the scores were transformed to approximate the scale of the raw counts. This was done by first multiplying the standardized scores by the standard deviation of the original raw counts, thereby changing the scale of the measures back to that of the raw counts. Then a constant was added so that the minimum value of the measures was equal to zero. In this way the measures could be treated as though they were raw, yearly counts.

REFERENCES

- Allison, Paul David
 • 1976 Processes of Stratification in Science.
 Ph.D. dissertation, Department of Sociology,
 University of Wisconsin, Madison.

- Astin, Alexander W.
1971 *Predicting Academic Performance in College*. New York: Free Press.
- Caplow, Theodore and Reece McGee
1958 *The Academic Marketplace*. New York: Doubleday.
- Cartter, Allan M.
1966 *An Assessment of Quality in Graduate Education*. Washington, D.C.: American Council on Education.
- Chemical Abstracts Service
1955- Chemical Abstracts. Columbus: American
1973 Chemical Society.
- Cole, Jonathan R. and Stephen Cole
1973 *Social Stratification in Science*. Chicago: University of Chicago Press.
- Crane, Diana
1965 "Scientists at major and minor universities: a study in productivity and recognition." *American Sociological Review* 30:699-714.
1970 "The academic marketplace revisited." *American Journal of Sociology* 75:953-64.
- Curtis, Richard B.
1969 *The Invisible University: Postdoctoral Education in the United States*. Washington, D.C.: National Academy of Sciences-National Research Council.
- Hagstrom, Warren O.
1967 *Competition and Teamwork in Science*. Final Report to the National Science Foundation for research grant GS-657. Department of Sociology, University of Wisconsin, Madison.
1968 "Departmental prestige and scientific productivity." Paper presented at the annual meeting of the American Sociological Association, Boston.
- Hagstrom, Warren O. and Lowell L. Hargens
1968 "Mobility theory in the sociology of science." Paper presented at the Cornell Conference on Human Mobility, Ithaca.
- Hargens, Lowell L.
1976 "A note on standardized coefficients as structural parameters." *Sociological Methods and Research* 5:247-56.
- Hargens, Lowell L. and Warren O. Hagstrom
1967 "Sponsored and contest mobility of American academic scientists." *Sociology of Education* 40:24-38.
- Institute for Scientific Information
1961- *Science Citation Index*. Philadelphia: Institute for Scientific Information.
- Long, J. Scott
1977a "Productivity and position in the early academic career." Ph.D. dissertation, Department of Sociology, Cornell University.
1977b "Entrance into the academic career." Unpublished paper. Department of Sociology, Cornell University.
- Manis, Jerome G.
1951 "Some academic influences upon publication productivity." *Social Forces* 29:267-72.
- Merton, Robert K.
1942 "Science and technology in a democratic order." *Journal of Legal and Political Sociology* 1:115-26.
1973 *The Sociology of Science*. Chicago: University of Chicago Press.
- Reskin, Barbara F.
1977 "Scientific productivity and the reward structure of science." *American Sociological Review* 42:491-504.
1978 "Academic sponsorship and scientists' careers." Unpublished paper. Department of Sociology, Indiana University.
- Roose, Kenneth D. and Charles J. Andersen
1970 *A Rating of Graduate Programs*. Washington, D.C.: American Council on Education.
- Schoenberg, Ronald
1977 "Dynamic models and cross-sectional data: estimation of dynamic parameter in cross-sectional data." *Social Science Research* 6:197-210.
- Weber, Max
[1919] "Science as a vocation." Pp. 129-56 in
1946 H. H. Gerth and C. W. Mills (eds.), *From Max Weber*. New York: Oxford University Press.
- Wilson, Logan
1942 *The Academic Man*. New York: Oxford University Press.
- Zuckerman, Harriet and Robert K. Merton
1971 "Patterns of evaluation in science: institutionalization, structure and functions of the referee system." *Minerva* 9:66-100.

OCCUPATIONAL SEX IDENTIFICATION AND THE ASSESSMENT OF MALE AND FEMALE EARNINGS INEQUALITY*

STEVEN D. McLAUGHLIN

Utah State University

American Sociological Review 1978, Vol. 43 (December):909-921

Recent research on sex differences in the process of earnings attainment has relied heavily on occupational prestige or status as a determinant of male and female earnings. Utilizing occupational level data from the U.S. Census and the Dictionary of Occupational Titles, this paper presents evidence suggesting that the use of prestige may lead to misspecification when the earnings attainment of males and females is compared. The data reveal differences in the nature of the occupational task by occupational sex identification within levels of prestige. Since some of these task differences are shown to be income relevant, the task-based earning potential of male dominated occupations is higher than the earning potential of equally prestigious occupations dominated by females. Finally, after the effects of both prestige and the nature of the task have been controlled, the sex identification of the occupation is shown to have a substantial impact on male and female earnings.

Among the several manifestations of gender stratification in the United States the most visible is the relative position of men and women in the labor force. In spite of dramatic increases in rates of female labor force participation, women continue to work in a relatively small number of sex-typed occupations for earnings considerably lower than their male counterparts. The perpetuation of occupational sex-typing and the male-female earnings differential as constants over a period of time in which remarkable changes in rates of female labor force participation were taking place provides evidence suggesting the existence of gender stratification in the labor force. The continuation of a stable system of occupational segregation (Hooks, 1947; Gross, 1968; Oppenheimer, 1968) is clearly related to the fact that women working full time and year-round earn about 60% of male earnings (U.S. Department of Labor, 1974).

The nature of the relationship between sex differences in occupational structure and sex differences in earnings is central to the process of gender stratification, yet this relationship is poorly understood.

There remains a great deal of uncertainty regarding the manner in which occupational characteristics influence the earnings attainment of individuals. Treiman and Terrell (1975a) have suggested that women earn less than men because they are relegated to low paying jobs while Hodge and Hodge (1965) have pointed out that the incumbents of some occupations may have lower earnings because the occupations are dominated by women. Polacheck (1975) suggests that women are clustered in low paying occupations because they experience a higher rate of skill atrophy as a function of their intermittent labor force participation. Others (e.g., Bergman, 1973) have argued that the crowding of women into a small number of occupations inflates the labor supply for those occupations, and, therefore, lowers the wage rates. On the other hand, research focusing on sex differences in the process of earnings attainment (e.g., Suter and Miller, 1973; Treiman and Terrell, 1975b) has deemphasized the occupational structure and has suggested that differences in the male and female rates of return for investments in human capital account for a large part of the earnings differential.

The diversity of these explanations of the earnings differential and its relationship to occupational sex-typing is due, in

* Address all communications to: Steven D. McLaughlin; Department of Sociology; Utah State University; Logan, UT 84322

part, to differences in levels of analysis. The research focusing on the impact of occupational structures on earnings takes the occupation as the unit of analysis while investigations emphasizing the process of income attainment via returns to human capital take the individual as the unit of analysis. The research conducted within either of these two levels has implications for research conducted at the other level, the *ecological fallacy* notwithstanding. Specifically, the individual-level research has not been sensitive to the implications of occupational structure on the differential earnings attainment of males and females.

This paper addresses one of the problems faced by the individual-level research which is generated by this lack of attention to the impact of occupational structure. The argument will be developed in the following sections by: (1) suggesting that occupational-level research findings pose a potential threat to the validity of the findings from the individual-level attainment research, and (2) examining occupational-level data to determine the extent to which the earnings attainment literature may be biased by the absence of attention to occupational structure.

Sex Difference in the Earnings Attainment Process

The documentation of sex differences in the process by which males and females acquire earnings has been conducted almost exclusively at the individual level of analysis. The dominant approach to the problem has been to regress separately male and female earnings on education, experience (if available), extent of labor force participation (hours worked per year), and occupational status or prestige.¹ Thus, the only empirical recognition given to the fact that men and

women work in occupations with potentially different structures has been the inclusion of an index of occupational prestige in the equation. The central thrust of this research has been to compare the size of the regression coefficients from the male and female equations in order to determine if the sexes experience similar rates of earnings return for their education, experience, hours worked, and occupational prestige.

The findings have been remarkably consistent. Suter and Miller (1973), Treiman and Terrell (1975b), and Featherman and Hauser (1976) all have documented sex differences in rates of earnings returned in exchange for education and occupational prestige. Among demographically equivalent groups of males and females working in equivalent occupational prestige categories, females experience less earnings in return for their education than males. Furthermore, among demographically equivalent males and females with similar educational credentials, males enjoy a higher rate of earnings return for their occupational prestige.

These findings are suggestive of economic discrimination in that male and female workers appear to be paid at different rates in exchange for the same credentials. However, in order to make this interpretation, it is necessary to have confidence in the specification of the earnings equations. More specifically, a discrimination interpretation of these findings assumes that there are no other variables with a unique earnings effect which vary systematically by sex. If such a variable is found the equations are said to be misspecified and comparisons of the male and female coefficients may be misleading.

One potential threat to a discrimination interpretation of these individual-level findings may be found in differences that exist in the nature of the occupations within which males and females typically work. Although the individual-level equation includes occupational prestige as a determinant of earnings, there may be occupational characteristics which vary within categories of prestige. Since prestige is not intended to measure all of the income-relevant dimensions of an oc-

¹ The use of an index of socioeconomic status vs. prestige for the comparison of male and female income attainment has been discussed by Treiman and Terrell (1975b) and McClendon (1976). Although there are clear conceptual differences between the two variables (Hauser and Featherman, 1977), both prestige and status have been utilized in previous research with only minor differences in the findings.

cupation, it is likely that some of the within prestige level variance in occupational structure is associated with earnings. This in itself is not a problem. However, if these within-prestige differences in occupational structure vary with the sex identification of the occupation, a potential misspecification emerges. The earnings structure of a male dominated occupation may be quite different from the earnings structure of a female dominated occupation even though they are identical in prestige. Given the extent to which the sexes are occupationally segregated, it is reasonable to assert that most female respondents in the individual-level earnings attainment research are assigned prestige scores from female dominated occupations. Consequently, the income-relevant dimensions of their occupations may differ from the income-relevant dimensions of the male workers with whom they are being compared. Thus, the consistently documented differences in rates of return realized by the sexes for education and occupational prestige may be partially due to the income-relevant characteristics of their occupations which are not measured with prestige. These occupational-level issues and the resulting implications for individual-level research form the central focus of this research.

*Occupational Characteristics,
Occupational Sex-Typing, and the
Earnings Attainment Process*

Although a direct examination of the implications of occupational structure for individual-level analyses of sex differentials in earnings does not exist in the occupational-level literature, there are investigations which suggest that there may be income-relevant differences in the nature of occupations dominated by men as opposed to those dominated by women. Oppenheimer's (1968) analysis of female dominated occupations revealed that these occupations tend to have incumbents with higher median education than the labor force at large. In addition, she found that earnings were lower for both men and women who worked in these occupations. Oppenheimer's analysis suggests that there may be differences in

the characteristics of female dominated occupations that result in lower earnings for incumbents regardless of their sex. This does not imply that all or even a major part of the earnings gap is due to the nature of the occupation. Males and females may still be paid at different rates for equivalent work, but a portion of the earnings gap may be due to occupational characteristics. The question as to how much of the earning differential is due to occupational structure remains an empirically testable issue.

Treiman and Terrell's (1975a) occupational-level analysis also provides evidence of the lower earnings realized by both male and female incumbents of female dominated jobs. They found that although women universally were paid less than men, the jobs that paid men well also paid women well. However, women were found to dominate the occupations which combined two characteristics: a high demand for educational attainment coupled with low earnings. Once again these findings are inconclusive. The low rate of pay found in female dominated jobs could be due to individual-level phenomena (lower rates of return realized by women), a tendency to pay incumbents of female dominated jobs at a lower rate simply because the occupation is sex-typed, or the female dominated occupations may actually differ from other occupations in the income-relevant characteristics of the task. It is this last alternative explanation that, if substantiated, poses a major threat to the validity of the previously cited individual-level research. If there are income-relevant differences in the nature of the occupational task that are not measured with occupational prestige, then the comparisons of male and female rates of return to education and prestige may be biased.

In order to assess these issues, the following hypotheses stated as questions will be examined: (1) Within levels of occupational prestige, is the nature of the task performed different for male dominated occupations than for female dominated occupations? (2) Do the income-relevant task characteristics of male and female dominated occupations differ within categories of occupational prestige? (3)

Does the sex identification of the occupation influence earnings within occupations involving similar tasks and with equal prestige?

Data and Methods

The data were drawn from two separate sources: the *Dictionary of Occupational Titles* (U.S. Department of Labor, 1965) and published 1970 census data. The *Dictionary of Occupational Titles* (DOT) was employed as a source of information concerning the nature of the occupational task and the requirements made on the occupation's incumbents. The 1970 census was utilized as the source for the number of male and female incumbents and the median earnings of each occupation.

The basic coding scheme used to classify occupations in the DOT was based on the assumption that all jobs involved some relationship to people, to data, and to things. It was assumed that these relationships were continuous such that job functions proceeded from the simple to the complex with regard to each relationship. Taken as a set of indicators, these three variables (classifications regarding data, people, and things) represented a description of each job's content. (For a description of the history and rationale of the DOT job evaluation system, see Fine and Heinz, 1958.)

In spite of an attempt to create continuous indicators of the data, people, and things dimensions, only the complexity of the job's relationship to data seems to approach this objective. (See Appendix A and B of Volume 1 of the DOT for a detailed description of all occupational characteristic indicators.) In light of the roughness of the DOT measures, "data" were trichotomized to form a scale with the following levels: (1) the lowest level in which there was no relationship to data at all or the worker needed only to compare readily observable characteristics; (2) the middle range of complexity in which the worker performed operations on data such as computing, compiling, and analyzing; and (3) the highest level in which data was coordinated for decision making or was integrated for the discovery of knowledge. Since the ordering of complexity levels for

the people and things dimensions was considerably less clear, both were dichotomized to form dummy variables to represent the presence or absence of an occupational relationship to people or things.

A problem raised by the use of these DOT characteristics stems from recent charges that the requirements of female dominated occupations were undervalued. Although this potential bias cannot be ignored it will be assumed that the job evaluations were valid enough to permit the rough trichotomy of "data" and the two dichotomies of *people* and *things* described above.

In addition to these baseline data, the third edition of the DOT incorporates detailed worker traits measured from a sample of 4,000 occupations reported in the DOT (U.S. Department of Labor, 1956). Among the occupational characteristics included are general and specific training required as well as aptitudes and working conditions which characterized the worker-job relationship.

General educational development (GED) was measured with a six-point scale reflecting the amount of formal and informal training necessary for average job performance. It was intended to tap the dimensions of education relevant to the development of reasoning ability, communication skills, and mathematics. The second training requirement, specific vocational preparation (SVP), was operationalized with a nine-point scale representing the increasing amount of time required to learn the techniques of a specific occupation. It included vocational education, apprentice work, and on-the-job training.

The aptitudes recorded for each occupation were the following: intelligence, verbal skill, numerical skill, spatial comprehension ability, form perception, clerical perception, motor coordination, finger dexterity, manual skill, eye-hand-foot coordination, and color discrimination. Each of these aptitudes was assigned to an occupation on the basis of a one- to five-point scale indicating increasingly higher requirements for that aptitude.

The DOT's measures of the job's working conditions included a five-point scale

of physical strength required, the presence or absence of a need to climb, stoop, talk, see, or handle objects. It also included indicators of the presence or absence of a requirement for the worker to experience extremes of hot and cold, loud noises, fumes, and of the job's requirement to work inside, outside, or both.

In addition to this extensive list of occupational characteristics, it was necessary to obtain additional information regarding the earnings and the sex identification of the occupation. Since the DOT contains no information regarding either earnings or the number of male and female incumbents, the 1970 published census data were utilized as a source (U.S. Bureau of the Census, 1973). In order to merge the two separate data sets, it was necessary to match each 1970 census three-digit occupational code to a DOT occupational title. Although there was a close correspondence between the census classifications at the professional level, the DOT was considerably more detailed than the census for managerial, sales, crafts, and operative occupations. This lack of correspondence required a careful reading of each DOT job description and a concurrent examination of the "Alphabetical Index of Occupations and Industries" (U.S. Bureau of the Census, 1970) to determine the best possible match among several DOT titles for each census title. In a few cases, the census title was so general that an accurate match could not be provided (e.g., "laborers not elsewhere classified"). In these cases the occupations were dropped from the analysis.²

This matching procedure resulted in a data set containing the DOT occupational characteristic information along

with the number of male and female incumbents who were in the experienced civilian labor force in 1970 and the median earnings of the experienced civilian labor force who worked 50-52 weeks in 1969 for each of 378 detailed 1970 occupational titles.

The final piece of information required for this research was an index of occupational prestige. Prestige was selected rather than an index of socioeconomic status because of the problems inherent in addressing issues of sex differences with an index based on male income and male education (Duncan, 1961). Moreover, Bose (1973) has shown that socioeconomic status indicators are more responsive to the effect of sex on the incumbent than are prestige indicators. (See Parnes et al., 1970; Heyns and Gray, 1973; and Haug, 1975, for further evidence on this issue.) Consequently, prestige scores from the adaptation of the Hodge-Siegel-Rossi index for 1970 census occupations were assigned to each occupation (Van Dusen and Zill, 1975).

With the assembly of occupational data on earnings, numbers of male and female workers, occupational prestige, and the set of DOT occupational characteristics, it became possible to assess empirically the extent to which occupational characteristics vary systematically by the sex identification of the occupation within categories of occupational prestige. Moreover, those occupational characteristics (if any) that differentiate prestige-equivalent male and female occupations could be examined for their relevance to income attainment.

The Occupational Characteristics

The first step in conducting the analysis was to construct measures of the occupational task from the set of variables drawn from the DOT. Initially, the variables which described the physical demands and working conditions were separated from the variables which tapped a skill-requirement dimension of the task. The indicators of the presence or absence of a need to climb, stoop, see, or to handle objects, along with the presence or

² Recent work by Temme (1975) and Spenner and Temme (1977) has resulted in the merging of 595 census occupation-industry classifications with the DOT classification system. Their procedures were more efficient than those described here since they had access to two Current Population Survey data sets containing both census and DOT classifications for all sampled members of the labor force. Thus, they were able to empirically weight each DOT characteristic according to its representation in the labor force. Unfortunately, the current Spenner-Temme merging of the DOT with census categories does not contain a majority of the DOT variables utilized in this research.

absence of loud noises, temperature extremes, and fumes were eliminated from the analysis due to the fact that these measures of working conditions were very nearly constant. The presence or absence of a need to talk was included with the skill-related indicators due to its correspondence with a verbal skill requirement. The remaining non-skill-related variable, physical strength, was retained in the analysis because of its probable relationship to both the occupation's sex ratio and the occupation's wage structure. Since it was not seen as a skill indicator, it was retained as a separate variable, independent of the set of skill-related indicators.

The DOT's basic data, people, things classification, along with the aptitude and training requirement measures were selected as indicators of the job's skill requirements. Since many of these variables were drawn from the same domain of content (e.g., the *talking requirement* and *verbal skill*) and consequently were highly correlated, it was necessary to create composite indicators. This was accomplished with a series of exploratory principle factor, orthogonal factor analyses with varimax rotation. In order to create composite indicators that had both a clear conceptual interpretation and were highly consistent with the data, all variables demonstrating a communality of less than .50 were dropped one at a time from the analysis. This procedure resulted in the elimination of spatial comprehension ability, clerical skill, eye-hand-foot coordination and color discrimination. The correlation matrix with means and standard deviations for the included variables are given in Table 1. The factor structure and communalities are displayed in Table 2.

An inspection of the factor loadings revealed a remarkably "clean" solution in that each variable was clearly loaded on one factor and the set of variables loading heavily on each factor seemed to share a common conceptual theme. Although this factor analysis was exploratory, the results communicated a great deal of face validity and thus it was with some confidence that the three resultant factors were utilized as indicators of the nature of the occupational task. Factor I was labeled as

cognitive skill because of the high loadings from data, the two measures of training level required, and the aptitudes of intelligence, verbal skill, and numerical skill. Factor II was labeled *manipulative skill* because of the high loadings from the four aptitudes relevant to skillful physical manipulation: form perception, motor coordination, finger dexterity, and manual skill. Factor III was identified as a *social skill* dimension due to its high loadings from the *people* and *talking* variables and the high negative loading from the variables assessing the complexity of the job's relationship to things. Linear composite indicators of each of these factors were then created with factor score coefficients generated with a regression technique (Harmon, 1967; Alwin, 1973). These coefficients, when used in a standardized regression equation, form the best possible least-squares estimation of the factors. With the single physical strength indicator these three factors complete the set of four variables measuring the nature of the occupational task.

The sex identification of each occupation was determined by the ratio of male to female workers within each occupational title in 1970 (U.S. Bureau of the Census, 1973). Unfortunately, no consensus exists regarding the operational rules for establishing *male* and *female* occupational categories. Following the work of Jusenius (1975), the labeling of an occupation as typically male or female was based on the proportion of the 1970 labor force that was female (38.1%). Any occupation in which women made up 48.1% of the workers ($38.1 + 10$) or more was labeled as a *female occupation*. This group contained 82 occupations or 24% of the total set of occupations. All occupations in which female workers represented 28.1% ($38.1 - 10$) or less of the occupation's incumbents were labeled as *male occupations*. This occupational category contained 247 (65%) of the occupations. The remaining 49 (11%) occupations in which women represented between 28.2 and 48.0% of the workers were labeled *mixed occupations*.

An alternative methodological approach to the analysis of the sex identification of the occupation is the use of the occupa-

Table 1. Correlations, Means, and Standard Deviations of the Selected DOT Characteristics

	Data	People	Things	Talking	GED	SVP	Intelligence	Verbal Skill	Numerical Skill	Form Perception	Motor Coordination	Finger Dexterity	Manual Skill
Data	1.0												
People	.276	1.0											
Things	-.328	-.749	1.0										
Talking	.370	.686	-.633	1.0									
GED	.818	.292	-.357	.400	1.0								
SVP	.743	.125	-.083	.265	.816	1.0							
Intelligence	.728	.367	-.455	.437	.881	.710	1.0						
Verbal skill	.808	.415	-.513	.488	.875	.674	.921	1.0					
Numerical skill	.767	.223	-.326	.365	.855	.727	.875	.869	1.0				
Form perception	.395	-.150	.232	-.056	.469	.512	.432	.390	.507	1.0			
Motor coordination	.109	-.188	.329	-.227	.168	.250	.136	.092	.162	.636	1.0		
Finger dexterity	.271	-.070	.238	-.119	.351	.446	.296	.237	.355	.766	.860	1.0	
Manual skill	.102	-.197	.340	-.213	.176	.254	.100	.028	.177	.644	.684	.780	1.0
\bar{X}	2.13	.386	.558	.466	3.81	5.63	3.44	3.29	2.88	2.83	2.77	2.67	2.78
S.D.	.85	.49	.50	.50	1.29	2.11	.987	1.10	1.18	.78	.63	.63	.63

Table 2. Varimax Rotated Factor Matrix for Selected DOT Characteristics

	FACTOR I	FACTOR II	FACTOR III	
	Cognitive Skill	Manipulative Skill	Social Skill	Communality
Data	.825	.095	.181	.722
GED	.918	.162	.202	.909
SVP	.795	.257	.016	.699
Intelligence	.870	.124	.305	.865
Verbal skill	.883	.061	.355	.910
Numerical skill	.892	.164	.157	.847
Form perception	.445	.698	-.168	.714
Motor coordination	.067	.849	-.140	.746
Finger dexterity	.223	.966	-.036	.984
Manual skill	.074	.791	-.172	.660
People	.131	-.075	.883	.821
Things	-.260	.286	-.786	.767
Talking	.319	-.158	.669	.574

tional sex ratio as a continuous variable. The categorization of occupations into male, mixed, and female is preferred over this alternative in that it permits a clear statement of the differences in the nature of the occupational task by the *sex identification* of the occupation. This technique is subject to criticism since the choice of cutoff points may influence the findings. With this danger in mind, the cutoff points were selected such that they maximized the correct identification of male and female occupations within the constraints imposed by the distribution of the sexes among those occupations.

Occupational Sex Identification and the Nature of the Occupational Task

The results of four multiple-classification analyses utilizing this classification of occupational sex identification are presented in Table 3. This technique was selected in order to provide answers to the following questions: (1) Do the mean values of the four indicators of the nature of the occupational task (cognitive skill, manipulative skill, social skill, and physical strength) vary by the sex-typing of the occupation? That is, do male, female, and mixed occupations display differences in that nature of the tasks? (2) Are there systematic differences in the nature of the occupational task

among male vs. female vs. mixed occupations *within* categories of occupational prestige?

The answer to the first question can be found in the unadjusted deviations column of Table 3. These values are deviations from the grand mean (which is zero in all cases due to standardization) of each dimension of the occupational task by the sex identification of the occupation. With the exception of manipulative skill, the effect of occupational sex identification on the nature of the task is statistically significant at the .001 level.

Female occupations demonstrate a remarkably low requirement for cognitive skill while the mixed occupations have the largest demand for this work dimension. The jobs dominated by women appear to be those with less requirements for symbolic thinking or for the solving of intellectual problems. The mixed occupations in which women have made a limited entrance, however, have the highest demands for cognitive skill. Consistent with common stereotypes of male vs. female jobs, the female and mixed occupations have a much higher demand for social skill and a lower demand for physical strength than male dominated occupations.

The second major question addressed by the data in Table 3 is whether or not these systematic differences in the nature of the task remain after prestige is controlled. In other words, do the task dif-

Table 3. Multiple Classification Analyses of the Four Occupational Task Indicators by the Sex Identification of the Occupation

	Unadjusted Deviations from the Mean	Deviations from the Mean Adjusted for Prestige
Cognitive Skill grand mean=0		
Occupational identification		
male	.09	.07
mixed	.37	.03
female	-.50	-.24
Manipulative Skill grand mean=0		
Occupational identification		
male	.05	.05
mixed	-.02	-.09
female	-.14	-.09
Social Skill grand mean=0		
Occupational identification		
male	-.26	-.26
mixed	.50	.40
female	.47	.55
Strength grand mean=0		
Occupational identification		
male	.21	.22
mixed	-.41	-.22
female	-.39	-.54

ferences among occupations with different sex identifications exist within categories of occupational prestige?

With the addition of prestige as a covariate, the possibility of an interaction becomes a salient issue. Consequently, an analysis of covariance was conducted in order to determine if the relationship between the occupational characteristics and occupational sex identification interacted with prestige. No significant interactions were found.

The second column of Table 3 provides the deviations, adjusted for prestige, for each of the four dimensions of the occupational task. As with the unadjusted mean deviations, the manipulative-skill dimension is not related to occupational sex identification. The mean deviations for the

other three dimensions, however, are statistically significant at the .001 level.

Within levels of occupational prestige the job requirements for cognitive skill, social skill, and physical strength are monotonically related to the sex identification of the job. As one moves from male to mixed to female jobs within levels of prestige the requirements for cognitive skill and physical strength decrease and the requirement of social skill increases. One can only conclude that occupational tasks of male and female occupations are not equivalent within categories of prestige. Stated another way, the occupational tasks found within equally prestigious male and female occupations are not equivalent.

Occupational Earning Potential and the Occupational Task

In order to determine the income relevance of the measured dimensions of the occupational task the median earnings of all year-round workers in 1969 was utilized to create an earning potential that was a direct function of the occupational characteristics. This was accomplished by regressing median earnings on each of the four occupational characteristics. Although male and female workers may be rewarded differently for these occupational characteristics, we utilized total earnings in this analysis, rather than separately regressing male and female earnings on the characteristics. The issue at hand is whether or not the characteristics are income-relevant. The construction of this occupational characteristic-based indicator of earning potential specifically was designed to avoid the confounding effects of differing rates of return by sex. It also should be noted that this indicator of occupational earning potential ignores the possibility of differences in rates of return to the occupational characteristics among the categories of occupational sex-identification. It is feasible, for example, that both male and female workers are paid at a higher rate in exchange for cognitive skill if they work in a male as opposed to a female dominated occupation. This potential interaction is an important research issue but is ignored here in order to

focus on the more immediate question: To what extent does the earning potential of equally prestigious male and female jobs differ as a function of the nature of the task? Thus, the earning potential of interest here is the average earnings for both males and females across all occupations. The issue of the impact of sex identification on earnings, net of prestige and net of the nature of the occupational task will be addressed separately for males and females in the following section.

The results of regressing median earnings on each of the four occupational characteristics revealed that the coefficients for social skill and physical strength were not significantly different from zero. Consequently, they were dropped from the equation and the coefficients for cognitive and manipulative skill reestimated. This procedure resulted in the following equation:

$$I = 8458.2 + 2495.1(\text{CS}) + 366.4(\text{MS}),$$

where *I* is the median earnings of the occupation, *CS* is cognitive skill and *MS* is manipulative skill. The regression coefficients are in unstandardized form with median earnings measured in dollars. These two occupational characteristics account for 44% of the variation in earnings ($R^2 = .444$).

With this equation, the occupational earning potential which is associated with these task characteristics was estimated by substituting each occupation's score on *CS* and *MS* into the equation and calculating each occupation's value for earning potential. This linear composite variable was then utilized as a dependent variable in the multiple classification analysis that appears in Table 4.

Without adjusting for prestige, female dominated occupations have a task characteristic-based earning potential that is \$2,208 less than mixed occupations and \$1,550 less than male occupations. These differences are a function of the occupational task weighted by the strength of their unique contribution to median earnings.

When these mean deviations are adjusted for prestige, female occupations have an earning potential \$674 less than

Table 4. Multiple Classification Analysis of Occupational Earning Potential by the Sex Identification of the Occupation

Occupational Identification	Unadjusted Deviations from the Mean	Deviations from the Mean Adjusted for Prestige
male	251	204
mixed	909	38
female	-1299	-636

Grand Mean=8458.2

mixed occupations and \$840 less than male occupations. Both the adjusted and the unadjusted deviations are significant at the .001 level. Since there is a potential for an interaction between occupational sex identification and earning potential with prestige, an analysis of covariance was conducted to test for this effect. No significant interaction effects were found.

These data demonstrate that the earning potential of male, female, and mixed occupations varies as a result of their differing cognitive and manipulative skill requirements within levels of occupational prestige. Consequently, the earning potential (based on the nature of the occupational task) is not likely to be equivalent for men and women working in prestige equivalent jobs typical of their sex.

Occupational Sex Identification and Actual Earnings

To this point the analysis has considered only the income differences among male, female and mixed occupations that can be attributed to the two income-relevant occupational characteristics. We have shown that the tasks performed in equally prestigious occupations differ by sex identification and that the earning potential of the occupation is influenced by these task differences. These findings lead to an additional question: If the effects of both the nature of the task and prestige are controlled, do earnings differ by the sex identification of the occupation?

The analysis suggested by this question appears in Table 5. The top panel displays the results of a multiple classification analysis in which deviations from the grand mean of median year-round male

Table 5. Multiple Classification Analysis of Male and Female Earnings by Occupational Sex Identification

Male Earnings		
Occupational Identification	Unadjusted Deviations from the Mean	Deviations from the Mean Adjusted for Occupational Characteristics and Prestige
male	527	628
mixed	1135	-261
female	-1925	-1254

Grand Mean=9407.4

Female Earnings		
Occupational Identification	Unadjusted Deviations from the Mean	Deviations from the Mean Adjusted for Occupational Characteristics and Prestige
male	583	443
mixed	319	-382
female	-1524	-758

Grand Mean=6138.4

earnings by sex identification are shown in the first column. These deviations are simultaneously adjusted for the four task characteristics and prestige in the second column. A parallel analysis utilizing median female earnings is presented in the lower panel.

The well-established pattern of female dominated occupations having lower earnings is clear. Both males and females are rewarded for working in male dominated occupations and penalized for working in female dominated occupations. Male workers appear to gain considerable advantage from working in mixed occupations. However, as the adjusted mean column indicates, this advantage is due to the prestige and the nature of the occupational task.

The most striking result appears in the adjusted means column. Even when the occupational characteristics and prestige are controlled, the sex identification of the occupation has a significant effect on the earnings of both males and females. Men working in male dominated occupations enjoy a \$1,882 earnings advantage over

their male contemporaries working in female dominated occupations that are identical in terms of prestige and the four indicators of the occupational task. Females working in male dominated jobs experience a \$1,201 advantage over their female counterparts working in equivalent female dominated occupations. It is clear that the sex identification of an occupation has an impact on earnings which is not accounted for by either prestige or the four measured dimensions of the occupational tasks.

Conclusions

The initial thrust of this research was the establishment of occupational task differences by the sex identification of the occupation which exist within levels of prestige. The substantiation of these task differences can be taken as evidence that there are occupational dimensions not measured by prestige which vary by sex identification. Taken by itself, this finding is not surprising. However, when followed by evidence that some of these within-prestige task differences are related to earnings, these findings have implications for the individual-level income attainment research. Since labor force participation is occupationally sex segregated (82% of all female workers are in female occupations and 74% of all males are in male occupations), it is reasonable to suggest that, given the above findings, the earning potential of most male and female workers within equally prestigious occupations is not equivalent. When a male and a female work in different occupations of equal prestige, it is very likely that their task-based earning potentials are different. With this in mind, the findings of male-female earning attainment research must be given a second look. The well-documented finding that earnings returned in exchange for occupational prestige and education are different for the sexes may be due partially to the fact that the earning potential of prestige-equivalent occupations occupied by males and females is different. When occupational prestige is utilized in models of income attainment or as the dependent variable in models of

occupational attainment the intention is to measure the income in return for equivalent work or the attainment of equivalent occupations. The findings of this research, however, suggest that prestige does not provide an indicator of equivalent occupations when utilized in research comparing male and female dominated occupations. In fact, the earning potential of prestige-equivalent occupations differs such that males tend to be clustered in occupations with greater income potential even within categories of prestige. This is consistent with the notion that many men and women who attain equally prestigious occupations do not attain equal occupational earning potential.

When the subject of investigation is individual-level sex differences in the process of earning attainment, the use of prestige as the sole occupational indicator is inappropriate and may lead to misspecification. When the issue at hand is whether or not males and females experience "equal pay for equal work" a set of indicators which describes the nature of the occupational task is far more appropriate than the use of prestige alone.

The conceptualization of an occupation as a matrix of differentially evaluated characteristics rather than as prestige will better demonstrate the manner in which human capital investments such as education provide access to specific types of occupations. If occupations, conceived as a set of characteristics, are employed in models of earning attainment, the ability of individuals to exchange specific types of work for earnings can be more directly assessed. The conceptualization of occupations as a set of characteristics addressed the real issues of equal pay for equal work and equal opportunity for employment. Although good data on occupational characteristics are often difficult to obtain, future research on occupational and earning attainment will benefit from the use of occupational characteristics as more direct indicators of the income-relevant dimensions of the occupation.

The final analysis serves to demonstrate that although the nature of the task within equally prestigious occupations varies by sex identification, this does not account for the earning differentials among male,

mixed and female occupations. Holding constant the nature of the task as well as prestige, we find the earnings of both males and females is greater in male than in female dominated occupations. This earning differential may be a function of occupational characteristics not measured in this research or a function of institutionalized discrimination at the occupational level.

REFERENCES

- Alwin, Duane F.
1973 "The use of factor analysis in the construction of linear composites in social research." *Sociological Methods and Research* 2:191-214.
- Bergman, Barbara
1973 "The economics of women's liberation." *Challenge* 16:11-7.
- Blau, Peter M. and Carol L. Jusenius
1976 Economists' approaches to sex segregation in the labor market: an appraisal." Pp. 181-200 in Blaxall and Reagan (eds.), *Women and the Workplace*. Chicago: University of Chicago Press.
- Bose, Christine E.
1973 *Jobs and Gender: Sex and Occupational Prestige*. Baltimore: Johns Hopkins University Press.
- Duncan, Otis Dudley
1961 "A socioeconomic index for all occupations." Chap. 6 in A.J. Reiss (ed.), *Occupations and Social Status*. New York: Free Press.
- Featherman, David L. and Robert M. Hauser
1976 "Sexual inequalities and socioeconomic achievement in the U.S. 1962-1973." *American Sociological Review* 41:462-83.
- Fine, Sidney and C. A. Heinz
1958 "The functional occupational classification structure." *Personnel and Guidance Journal* 37:180-92.
- Gross, Edward
1968 "Plus ça change...? The sexual structure of occupations over time." *Social Problems* 16:198-208.
- Harmon, Harry H.
1967 *Modern Factor Analysis*. Chicago: University of Chicago Press.
- Haug, Marie R.
1975 "Sex role variations in occupational prestige ratings." *Sociological Focus* 8:47-56.
- Hauser, Robert M. and D. L. Featherman
1977 *The Process of Stratification*. New York: Academic Press.
- Heyns, B. and J. Gray
1973 "A socioeconomic index for women's occupations." Unpublished paper. Center for Educational Policy Research, Harvard University.
- Hodge, R. and P. Hodge
1965 "Occupational assimilation as a competi-

- tive process." *American Journal of Sociology* 61:249-64.
- Hodge, R., P. Siegel, and P. Rossi
1966 "Occupational prestige in the United States 1925-1963." Pp. 322-34 in R. Bendix and S. Lipset (eds.), *Class, Status and Power*. New York: Free Press.
- Hoffman, Lois W. and F. Ivan Nye
1975 *Working Mothers*. San Francisco: Jossey-Bass.
- Hooks, Janet M.
1947 "Women's occupations through seven decades." *Women's Bureau Bulletin* No. 218. Washington, D.C.: U.S. Government Printing Office.
- Jusenius, Carol L.
1975 "Occupational change, 1967-1971." Chap. 2 in U.S. Department of Labor, *Dual Careers*. Washington, D.C.: U.S. Government Printing Office.
- McClendon, McKee J.
1976 "The occupational status attainment process of males and females." *American Sociological Review* 41:52-61.
- Oppenheimer, V. K.
1968 "The sex labeling of jobs." *Industrial Relations* 7:219-34.
- Parnes, Herbert S., John R. Shea, Ruth S. Spitz, and Frederick A. Zeller.
1970 *Dual Careers*. Vol. 1. Washington, D.C.: U.S. Department of Labor, Manpower Research Monograph No. 21.
- Polacheck, Solomon W.
1975 "Potential biases in measuring male-female discrimination." *Journal of Human Resources* 10:205-29.
- Spenner, Kenneth I. and Lloyd V. Temme
1977 "Spenner-Temme occupation variables for the 1970 census occupation-industry classification." Mimeo. Boys' Town Center for the Study of Youth Development. Boys' Town.
- Suter, Larry E. and Herman P. Miller
1973 "Income differences between men and career women." *American Journal of Sociology* 78:962-74.
- Temme, Lloyd V.
1975 *Occupation: Meanings and Measures*. Washington, D.C.: Bureau of Social Science Research.
- Treiman, Donald J. and K. Terrell
1975a "Woman, work, and wages—trends in the female occupational structure." Chap. 7 in K. Land and S. Spilerman (eds.), *Social Indicator Models*. New York: Russell Sage.
1975b "Sex and the process of status attainment: a comparison of working women and men." *American Sociological Review* 40:174-200.
- U.S. Bureau of the Census
1970 *Alphabetical Index of Occupations and Industries*. Washington, D.C.: U.S. Government Printing Office.
1973 *Census of Population: 1970, Subject Reports, Final Report PC(2)-7A. Occupational Characteristics*. Washington, D.C.: U.S. Government Printing Office.
- U.S. Department of Labor
1956 *Estimates of Worker Traits Requirements for 4000 Jobs*. Washington, D.C.: U.S. Government Printing Office.
1965 *Dictionary of Occupational Titles*. Washington, D.C.: U.S. Government Printing Office.
1974 *Fact Sheet on the Earnings Gap*. Washington, D.C.: U.S. Government Printing office.
- Van Dusen, R. A. and N. Zill (eds.)
1975 *Basic Background Items for U.S. Household Surveys, Appendix C*. Washington, D.C.: Social Science Research Council.

COMMENTS

MARXIST AND "ORGANIZATIONAL" APPROACHES TO DELINQUENCY AND THE SOCIOLOGY OF LAW: CRUCIAL PROBLEMS IN "TESTING" THE PERSPECTIVES*

(COMMENT ON HAGAN AND LEON, ASR AUGUST, 1977)

Hagan and Leon (1977) attempt to test the Marxian conflict perspective by reexamining Anthony Platt's (1969; 1974) work on the invention of delinquency.¹ Although they address an important issue, their analysis is misleading on both empirical and theoretical grounds.

The first problem with Hagan and Leon's article is that their test of Platt's "Marxist" account of the origin of delinquency legislation and its subsequent ramifications is essentially inappropriate.² They claim to refute Platt's

conclusions based on his research in Chicago with their research effort in Toronto. The issue here is really one of external validity (i.e., are Platt's findings universal?), but Hagan and Leon appear to be testing internal validity (i.e., are Platt's conclusions for Chicago valid from the evidence?). This confusion between external and internal validity is a basic flaw.³ At best, they decrease the level of confirmation of Platt's theory, and, more importantly, their inattention to the data from Chicago is incongruous.

Secondly, Hagan and Leon (1977:589) castigate "the Marxian perspective" by asserting that the perspective is "(1) prone to logical errors, (2) largely unconfirmed, (3) often unconfirmable and (4) possibly quite frequently false"; however, those assertions are not supported adequately. Since the charge that the perspective is "possibly quite frequently false" will follow if the three earlier charges are correct, we will focus upon the preceding three charges. To note that the perspective is largely unconfirmed is to provide a comment on most of the perspectives, current and past, in sociology. A more serious charge is that the perspective is unconfirmable. If Hagan and Leon are correct, then by implication the Marxian perspective is discredited seriously; however, given the frequent lack of agreement re-

they are testing a Marxian perspective and *the* Marxian perspective. This will be addressed further later in our comment. More importantly however, they imply that the perspective will be proved either right or wrong based on their work. We doubt that their work will have that effect. That is, the perspective is neither right nor wrong. The hypotheses derived from the theoretical orientation are in a state of confirmation or disconfirmation. In one of the more curious statements in the entire article (p. 589) the authors indicate that the Marxian perspective has a "scientific base." In light of their comments regarding the weaknesses in this perspective, we would wonder what they consider that base to be.

³ Although Hagan and Leon do take issue with one assertion from Platt's (1974) work, the bulk of their concern is with the contents of *The Child Savers*. In this book, Platt makes it clear that his work is a case study. It is, he says, a "study which addresses the origins, composition and achievements of the child saving movement in the United States" (Platt, 1969:3). This is further stressed when he notes, "I have chosen to focus specifically on the child saving movement in Chicago . . ." (Platt, 1969:9).

* Address all communications to: Pat Lauderdale; Department of Sociology; 1114 Social Sciences; 267 19th Ave. S.; Minneapolis, MN 55455.

We would like to thank Jack Clark, Lyle Hallowell, Kathryn Harwig, Charlie Lakin, Charles Reasons, Jeff Stitt, and Stan Wheeler for a number of helpful suggestions. We are especially indebted to James Inverarity for his extensive critique of this paper.

¹ A few months before Hagan and Leon's research appeared, the ASR printed three comments (Greenberg, 1977; Hopkins, 1977; Reasons, 1977) which basically confronted another attempt to test a form of the conflict perspective (Chiricos and Waldo, 1975). The comments were directed toward theoretical, methodological, empirical and ideological issues in Chiricos and Waldo's (1975) article; however, the main thrust of all three comments was to emphasize that the conflict position is not necessarily wrong. It is somewhat ironic (cf. Matza, 1969) that most work on the so-called conflict perspective appears outside the ASA journals, that most of the tests of the perspective occur within the journals, and the comments arise shortly thereafter. At first glance it appears that the problem is one of outsiders vs. insiders; however, we will attempt to illustrate how this dichotomy is as false as the functional/conflict one. We hope that what has been said up to this point will help prevent this short piece from becoming another part of the "comment" cycle. Our criticisms of Hagan and Leon's work are not intended as "bad news" (cf., Chiricos and Waldo, 1977), but, instead, as an initial step at refining delinquency and sociology of law research.

² Hagan and Leon vacillate between saying that

garding the nature of evidence, it should not be surprising that we also find little consensus regarding its sufficiency.⁴ The charge that the perspective is unconfirmable raises the thorny issue of what are considered testable hypotheses within the perspective. Rather than presenting evidence on this issue, the authors simply ignore it.

Now, consider the charge that the perspective is prone to logical errors. Hagan and Leon argue that it is the perspective which is prone to the production of logical errors, notably those of argumentum ad hominem and affirmation of the consequences. While their charge is lodged against the perspective, they fail to specify anything in the perspective which makes it prone to the production of these errors. Their evidence on this point consists of briefly cataloguing various authors who they feel have erred. This approach comes dangerously close to committing the ad hominem fallacy which the authors claim to deplore.

Also, we think that the authors protest too much when they rely upon terminology that is not amenable to test (i.e., prone, largely, often, and possibly quite frequently). Since Hagan and Leon maintain they are interested in letting the evidence provide answers to the hypotheses, then they are also obliged to present testable statements (Stinchcombe, 1976).

Thirdly, we find the conclusions which Hagan and Leon draw from their retest of Platt's hypotheses one-sided. That is, the authors basically reject the bulk of Platt's research without paying sufficient attention to those situations where the evidence is ambivalent or supportive of Platt's arguments. For example, Hagan and Leon (1977:595) state,

J. J. Kelso, a moral entrepreneur who rose from news reporter to a high position in the child-care bureaucracy, collaborated with W. L. Scott, a professional counterpart and philanthropist who occupied an administrative position in the Ontario Supreme Court, to engineer a legislative movement whose organizational goal became the prevention of delinquency through juvenile probation work.

They go on to note that it is not clear "whether the eventual success of advocates of probation served the basic interests of the ruling elite

..." (1977:595). This conclusion is inappropriate, especially when one considers that at the outset of their work Hagan and Leon (1977:590) stressed,

Summarizing, Platt's primary concerns are that links between middle-class reformers and upper-class sponsors resulted in the wealthy and powerful using the passage of American delinquency legislation to invent new forms of youthful misbehavior and increase imprisonment, all in the larger interest of developing a specialized labor market and industrial discipline under corporate capitalism.

Hagan and Leon's research simply supports some of Platt's arguments (e.g., the social position of the Canadian reformers and sponsors; the development of a specialized labor market—probation officers),⁵ and their attempt to negate the Marxian perspective here does a disservice to their research and Platt's.

Fourthly, the data that Hagan and Leon present are confusing and in a number of cases anecdotal. The only presentation of a systematic series of data (Appendix I) is from 1912 to 1951. We find that the raw data are difficult to interpret. Obviously the number of youths in the eligible age range increased dramatically between 1912 and 1952. The use of a rate therefore would have made their data more intelligible and meaningful. We also suspect that it may have strengthened their position. The au-

⁵ It seems clear to us that both Platt and Hagan and Leon are using *specialized labor market* to refer to the development of a pool of individuals trained and otherwise suited to engage in industrial labor. However, we would argue that the development and growth of the occupation of probation worker is indeed an example of the growth of a specialized labor market which in fact came about as a result of the juvenile court movement, both in Canada and in the United States. Compare, for instance, the data which Hagan and Leon present regarding the growth of the probation department, with the assertion by Platt (1974:371): "Social work and professional child-saving provided new opportunities for career minded women who found the traditional professions dominated and controlled by men. These child-savers were members of the emerging bourgeoisie created by the new industrial order." There seems to us no incompatibility here, yet Hagan and Leon chose to ignore this particular passage. Although we have no data on this point, we feel that Hagan and Leon's dismissal of Platt's *industrial discipline* argument is premature. Writing in the United States during the middle of the time period considered by Hagan and Leon, one federal judge characterized probation as a "string around the neck of each probationer." Further, he offered the opinion that "no idle man will keep probation," and therefore required each probationer to sign a pledge that he would "work regularly" (McClintic, 1937:17-25).

⁴ Hagan and Leon note (1977:591, 595) that the correspondence and documents which they analyzed fail to show any evidence of interest being expressed by industrial elites. We doubt this "evidence" would satisfy the several authors castigated in the first part of their article. Furthermore, Marxists (e.g., Balbus, 1977) have in fact provided empirical demonstration of the impact of class interests on the legal system which would speak to the criticisms raised by Hagan and Leon.

thors note that 123 juveniles were institutionalized in the year preceding 1912; however, it is not clear why only these data are available or presented. Also, while they claim that 71 juveniles were sent to industrial schools in the first year of court proceedings (again, 1912), the entry in the table is 85. The authors' suggestion that *Suspended Sentences and Adjudgments* includes commitments to industrial schools, Working Boys Home, Orilla Hospital, Ontario Hospital and training schools is confusing (fn. d). Four of the totals for Court Dispositions are incorrect, one is remarkably in error (the 1932 total should be 1,044 not 1,483, assuming the figures in the columns are not the error).⁶ The increase in *Occurrences* (under probation department statistics) is striking. Those data are especially relevant to the discussion concerning the development of a specialized labor market; however, the authors do not discuss whether this data might speak to the issue of whether institutionalized control may have increased. They simply note that the official occurrences decreased from 1912 to 1952 while the number of probation officers increased from four to 14 (Appendix I indicates five to 15) and the number of occurrences handled unofficially by the probation officers increased approximately 900%. The important concept of institutionalization and its potential indicators is virtually ignored.⁷

Furthermore, in the interval from 1942 to 1944 the number of judges increased from two to five, while the number of juveniles institutionalized (using Hagan and Leon's definition) rose from 85 to 120 (the largest number of the entire period). The shifts in most of the rows from 1928 to 1930 and 1932 are dramatic; however, the authors do not provide any discussion or explanation of these and other variations. Using the data from Appendix I, one can compute a statistic, *Percent of Court Referrals Sentenced to Institutions*. This reveals that the percent of court referrals who were subsequently institutionalized declined consistently from 4.8% in 1912 to 1.5% in 1926. It then increased slowly but consistently to reach 5.0 in 1942, at which point it rose rather

dramatically to nearly 10% (9.8) for 1944. It remained near the 10% mark for the duration of the period included in the Appendix, ending in 1952 at nearly 12% (11.7). Explanation of these variations may be most instructive. Finally, the curvilinear shape of *Percent Referred to Court* and *Suspended Sentence and Adjudgments* is left unexplored. It is not clear how the row including adjournments differs from the one labeled *Dismissed and Withdrawn*.

Fifthly, the direct link that Hagan and Leon draw between conflict theory and the Marxist perspective is too simplified, as are many of their interpretations of Marxist concepts. While Hagan and Leon continually refer to the Marxist perspective, we suggest that such a perspective is not the monolith which they imply. Rather, it is characterized by significant differences and diverse view points (cf., Gold et al., 1975). Hagan and Leon are concerned primarily with instrumental Marxism which traces lines among individual actors as a means of demonstrating that legal institutions serve the ruling class. Alternatively, a Marxist structuralist perspective might suggest that those policy makers are shaped by economic conditions which put the state in the position of trying to deal with contradictions which continually become evident within the system. Therefore, they can "explain" why many governmental decisions appear to go against the capitalist class. From this perspective the actions of policy makers maintain the economic structure within which those very interests can be expressed.

Hagan and Leon's exclusion of the structuralist perspective (as well as the Hegelian-Marxist perspective) is particularly confusing, especially considering their claim in the abstract: "The emergence of the Marxian perspective and the logic of its premises are reviewed" (1977:587).

Finally, Hagan and Leon equate incarceration (institutionalized juveniles) with social control. This ignores the interpretation of Platt and most Marxist sociologists that social control includes the activities of institutionalized probation workers.

In summary, we maintain that important empirical findings and interpretations, in "Rediscovering Delinquency" have crucial problems. The data we have examined from the article are either irrelevant to the claims being made or support other conclusions. In addition, the theoretical work is misguided. The article does not help resolve some of the current debates over competing perspectives nor does it offer a coherent alternative to the perspective it claims to test. Unfortunately, the article may help perpetuate the overly simplis-

⁶ We assume that fn. d was meant to be attached to the row labeled *Institutionalized* rather than to *Suspended Sentence and Adjudgments*, and that the adding errors are editorial mistakes; however, we need further clarification from the authors.

⁷ We think it is a mistake for Hagan and Leon to argue that probation is not a formal, explicit system of coercion. Since juveniles may be placed on probation without having committed an actual criminal offense, the conclusion of the authors that it is an informal system is somewhat surprising.

tic debate between dichotomized perspectives (e.g., order vs. conflict; cf., Giddens, 1976).⁸

Pat Lauderdale
Gerald Larson
University of Minnesota

REFERENCES

- Balbus, Issac D.
1977 *The Dialectics of Legal Repression*. New Brunswick: Transaction.
- Chiricos, Theodore G. and Gordon P. Waldo
1975 "Socioeconomic status and criminal sentencing: an empirical assessment of a conflict proposition." *American Sociological Review* 40:753-72.
1977 "Reply to Greenberg, Hopkins, and Reasons." *American Sociological Review* 42:181-5.
- Giddens, Anthony
1976 "Classical social theory and the origins of modern sociology." *American Journal of Sociology* 81:703-29.
- Gold, David, Clarence Lo and Erik Wright
1975 "Recent developments in Marxist theories." *Monthly Review* 5:29-43; 6:36-41.
- Greenberg, David F.
1977 "Comment on Chiricos and Waldo, ASR December, 1975." *American Sociological Review* 42:174-5.
- Hagan, John and Jeffrey Leon
1977 "Rediscovering delinquency: social history, political ideology, and the sociology of law." *American Sociological Review* 42:587-98.
- Hopkins, Andrew
1977 "Comment on Chiricos and Waldo, ASR December, 1975." *American Sociological Review* 42:176-7.
- McClintic, George H.
1937 "Probation." *Federal Probation* 1:17-25.
- Matza, David
1969 *Becoming Deviant*. Englewood Cliffs: Prentice-Hall.
- Platt, Anthony
1969 *The Child Savers*. Chicago: University of Chicago Press.
1974 "The triumph of benevolence: the origins of the juvenile justice system in the United States." Pp. 356-89 in Richard Quinney (ed.), *Criminal Justice in America*. Boston: Little, Brown.
- Reasons, Charles E.
1977 "Comment on Chiricos and Waldo, ASR

- December, 1975." *American Sociological Review* 42:177-80.
- Stinchcombe, Arthur
1976 "Marxist theories of power and empirical research." Pp. 301-14 in Lewis Coser (ed.), *The Uses of Controversy in Sociology*. New York: Free Press.

CRASS CONFLICT: A REPLY TO LAUDERDALE AND LARSON*

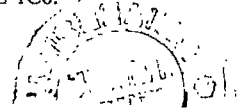
We regret that Lauderdale and Larson misunderstand empirical and theoretical aspects of our work (Hagan and Leon, 1977a), but we are grateful for the opportunity to speak to several of the issues they raise. Following a brief summary of the central points of our paper, we respond to their comments, in the order they are presented.

Our paper began with a review of the conflict-consensus debate and the Marxian perspective on crime and law that has emerged. We noted that while this work stemmed from an interest by Chambliss, Quinney, Platt and others in the influences of *various* class and status groups, more recently it has assumed a single-minded Marxian focus on class conflict and coercion. This shift is particularly conspicuous in a comparison of Platt's early Chicago study (1969) of *The Child Savers* and his later reinterpretation (1974) of his initial findings in an article, "The Triumph of Benevolence: The Origins of the Juvenile Justice System in the United States." The latter paper includes what we believe to be an overemphasis by Platt on the role of penal institutions, while it neglects the impact of probation as the distinctive disposition associated with the emergence of juvenile court legislation. Most significantly, Platt (1974:377) relies on an assumed increase in the use of penal institutions as the basis for concluding that "the juvenile court system was part of a general movement directed towards developing a specialized labor market and industrial discipline under corporate capitalism by creating new programs of adjudication and control. . . ." In this 1974 article, Platt does not restrict his conclusions to Chicago, but generalizes to the American system of juvenile justice.

In contrast, we found that in Toronto, passage of Canadian delinquency legislation had no immediate or dramatic effect on the use of penal dispositions. The numbers of juveniles institutionalized initially declined, and increased only later in rather modest amounts. On the other hand, the use of probation as a

⁸ Sociologists should be in a position to examine competing perspectives; however, we think most of the ubiquitous tests of one perspective vs. another in the areas of deviance and sociology of law are premature and do not help resolve present day dilemmas in sociology. The setting of one perspective vs. another may be a reasonable orientating strategy; however, it is not the final solution, especially when the data (regardless of the methodology) do not fit easily into any of the chic frameworks.

* Address all communications to: John Hagan; Erindale College; University of Toronto; Mississauga, Ontario L5L 1C6.



disposition and the engagement of probation personnel increased dramatically. These organizational developments had much more to do with the emergence and evolution of the Canadian juvenile court system than did the potential expansion of penal institutions and the use of institutional dispositions.

In their comments on our paper, Lauderdale and Larson first are concerned that we have confused the external and internal validity of Platt's findings. Their apparent preference is for a reexamination of the Chicago data in order to test the internal validity of Platt's original work. However, it is Platt, if anyone, who confuses the issue of internal and external validity. As indicated above, Platt's 1974 conclusions are not restricted to Chicago, even though no new systematic source of data is introduced.¹ Our interest was in broadening the sample to include at least a second city (Toronto) and country (Canada) in which to assess the generality of Platt's assertions. We assume that the rather different findings for Toronto encourage a reexamination of the internal validity of Platt's work, and we cited an American critique (Schultz, 1973) of Platt's premises that further suggests the need for such a reassessment. Nonetheless, the conclusion to our paper contents itself with a clear statement about external validity:

... this paper provides a considerably different account of the origin of delinquency legislation than is offered by Platt, and it is an account which encourages a reconsideration of the conditions under which this version of a Marxian, class conflict perspective can be usefully applied. (p. 597)²

Lauderdale and Larson are concerned secondly with our conclusion that the version of the Marxian perspective we examined is "(1) prone to logical errors, (2) largely unconfirmed, (3) often unconfirmable and (4) possibly quite frequently false" (p. 589). They argue that "[t]he charge that the perspective is unconfirmable raises the thorny issue of what are considered *testable* hypotheses within the perspective. Rather than presenting *evidence* on this issue, the authors simply ignore it" (emphasis added). However, the issue of what is testable is an issue of *logic*, not evidence, and we give considerable attention in our paper to the logical status of recent Marxian formulations. Lauderdale and Larson go on to complain that we "... fail to specify anything in

the perspective which makes it prone to the *production* of these errors" (emphasis added). Yet our conclusion, as stated, is that the *theorists*, not the theory, produced these logical errors. In fact, we argue that the underlying conflict-consensus debate can be very useful (see also Hagan, 1977; Hagan et al., 1977). This explains the use of the cautious phrasing ("prone, largely, often and possibly quite frequently") noted above. Lauderdale and Larson object that phrasing our assessment of Marxian formulations in such terms limits the possibilities of empirical testing. However, the problem lies with statements of the theory, not with our terminology. Thus we tested and found false Platt's assertion that penal dispositions increased with the passage of juvenile delinquency legislation. That few other of Platt's assertions were amenable to test is regrettable—and a comment on the state of the theory. It is as well a reason why we found it necessary to develop an organizational analysis of the emergence of Canadian delinquency legislation. However, for our critics to suggest that we "imply that the [Marxian] perspective will be proved either right or wrong based on . . . [our] work" is simply false.

Lauderdale and Larson's third point is their most crucial argument. They assert that

Hagan and Leon's research simply supports some of Platt's arguments (e.g., the social position of the Canadian reformers and sponsors; the development of a specialized labor market—probation officers), and their attempt to negate the Marxian perspective here does a disservice to their research and Platt's.

This summary of the situation is highly misleading, for it misrepresents both our arguments and Platt's. As we note in our paper, in 1973 Platt, rejecting his earlier conclusions, observed that "the problem with *The Child Savers* is that . . . [it] focuses too much attention on the middle-class reformers. . . ." Thus when Platt speaks in 1974 of a specialized labor market (see quote above), he clearly is not referring to the growth of the probation profession. Lauderdale and Larson concede this point, and deny their own argument, when they note in their footnote 5 that "[i]t seems clear to us that both Platt and Hagan and Leon are using *specialized labor market* to refer to the development of a pool of individuals trained and otherwise suited to engage in *industrial labor*" (emphasis added). Thus Platt's Marxian revision of his *Child Saving* views is explicitly concerned with the use of penal institutions, for youths, to instill industrial discipline, in preparation for industrial labor; Platt gives no attention here to the growth of probation as a specialized labor mar-

¹ Moreover, Platt's thesis frequently is referenced as an explanatory framework for the development of American, and Canadian, juvenile justice systems.

² On the importance of studying crime control strategies in different social, national and cultural contexts, see Hagan and Leon (1977b).

ket. Lauderdale and Larson's footnote indicates that they know this, but the text of their discussion proceeds as if they do not. Finally, Lauderdale and Larson's synthesis of our findings with Platt's perspective ignores one additional, and fundamental, difference. We found that there indeed was conflict about the passage of delinquency legislation in Canada. However, this conflict was not class-based. Rather, the conflict consisted of an ongoing debate between advocates of probation and representatives of the police who favored the more punitive policies emphasized by Platt; we found the former organizational group prevailing.

The fourth set of points made by Lauderdale and Larson involve the qualitative and quantitative data we present. The former are characterized as "anecdotal" and the latter as "systematic," yet "confusing." The glib characterization of the former data is, in our opinion, factitious. However, since Lauderdale and Larson say nothing more specific about this part of our analysis, we can say little more in its defense.³ We can clarify several points with regard to Appendix I, which will hopefully resolve the inconsistent description of this data as both systematic and confusing.

For example, Lauderdale and Larson are correct in suggesting that the arguments made in our paper would have been supported *even more strongly* if age based rates could have been calculated for the Appendix. That is, since the relevant age group increased dramatically between 1912 and 1952, and the absolute number of juveniles institutionalized showed no clear pattern of increase, the *proportion* of youths institutionalized must have dropped rather substantially. Unfortunately, the relevant Toronto census data necessary for these calculations do not exist for the early years of our analysis. Nonetheless, the raw data make our point clearly enough.

Closer reading of the text of our paper would have clarified several additional points of confusion in Lauderdale and Larson's reading of our Appendix. We note in the text (pp. 590-1) that the 1912 *Annual Report of the Juvenile Court* indicates that 123 juveniles were sent to industrial school *alone* in 1911. Seventy-one were sent to such schools in 1912, the first year in which the Juvenile Court formally operated and published official statistics. Additional data on other institutional dispositions in 1911 are not provided in the 1912 *Report*. However,

we do learn from this report that in 1912 another 14 youths were institutionalized in a working-boys home, two hospitals and/or a training school. This fact, included in the Appendix, yields the figure of 85 youths institutionalized. In the text we compare the 123 and 71 figures; using the 85 figure would also support our argument, but would be a less meaningful comparison. We make clear in the text that *institutionalization* refers to the use of industrial schools, a working boys home, two hospitals and a training school. However, during the typesetting, footnote d in Appendix I apparently was misplaced after *Suspended Sentences and Adjournments*—a combined category situated directly above *Institutionalization*. In any case, and as Lauderdale and Larson concede in their footnote 6, the text makes our categorization obvious.

Lauderdale and Larson further indicate that four of the column totals in the Appendix do not match the sum of the figures in the columns. The explanation for this is that five of 1,744 cases in column 1, one of 2,298 cases in column 4, five of 2,122 cases in column 7, and 439 of 1,483 cases in column 8 were not categorized by type of disposition. Only column 8 involves a sufficient number of cases to merit full discussion.⁴ These cases consist of youths placed on probation and under the care of a "big brother or sister," rather than a probation officer. Since these were probationary dispositions, and since they were grouped as such in preceding and succeeding years, we should have included them in the probation row. Including these cases results in an entry for the 1932 column of 538, a figure fully consistent with the balance of the data and with the arguments made in our paper.

Similarly, our text speaks of an increase from four to 14 probation officers from 1912 to 1952. There was as well a chief probation officer in the probation department in 1912, and a social investigator in 1952; this yielded an overall personnel increase from five to 15 employees. Footnote c to Appendix I indicates the more inclusive definition of probation officer used in constructing this part of the Appendix. Nevertheless, use of either set of figures, however defined, supports our argument of expansion in the probation bureaucracy.

Finally, Lauderdale and Larson are concerned that we did not (a) calculate a statistic,

³ Yet if it is the quantity of the qualitative data that is of concern, see Leon (1977) for a more detailed analysis of the sociolegal development of Canadian juvenile delinquency legislation.

⁴ The five cases in col. 1 could not be classified, even under the category *other*. The six cases in cols. 4 and 7 were apparently lost in the typing of the Appendix. As well, we should note that another entry apparently was reversed in the type-setting: the number of interviews in 1932 should read 9,948 instead of 9,984.

"percent of court referrals sentenced to institutions," (b) comment on the curvilinear shape of the percent of cases referred to court and the suspended sentence and adjournments figures, or (c) distinguish between adjournments and cases dismissed and withdrawn. The "percent of court referrals sentenced to institutions" statistic is rather misleading in that it omits consideration of two important trends noted earlier: that the number of court cases was declining during a period in which the adolescent population was increasing. Looking instead at the raw data, the absolute numbers of youths institutionalized, we are able to show the more significant fact that institutional populations remained rather strikingly stable over this 40-year period. The implication of this is that penal administrators were able to maintain the operating levels of their bureaucracies even though circumstances were changing rapidly around them. The additional curvilinear trends noted by Lauderdale and Larson were not central topics of our paper. Finally, the unique feature of an adjournment is that the case remains open to recall indefinitely, while a dismissal or a withdrawal terminates the proceedings in respect of the particular charge before the court.

Fifthly, Lauderdale and Larson suggest that we have constructed, by implication, a Marxian "monolith" which does not exist; rather, they argue, this perspective is characterized by "significant differences and diverse viewpoints." Our contention was simply that there was enough convergence between the work of Marxist criminologists at the time our article was written to engage in an empirical test of the propositions that this perspective offered about the origins of juvenile justice legislation. However, there is a danger, evidenced by Lauderdale and Larson's comment, that there will emerge so many "Marxisms" that almost any empirical finding will be consistent with some version of the "theory." The effect, of course, will be to make the "theory" untestable.

The last issue that Lauderdale and Larson raise may be the most significant for future research. Unfortunately, this issue is posed in a rather confusing way. Lauderdale and Larson assert that we "... equate incarceration (institutionalized juveniles) with social control. This ignores the interpretation of Platt and most Marxist sociologists that social control includes the activities of institutionalized probation officers." We do not deny that probation is a form of social control, and use of the term *institutionalized* in the same passage in two such different ways is unnecessarily confusing. The point made in our paper is that probation and prison are very different *types* of

social control. Probation involves a *cooptive* effort, including some restrictive provisions, to reintegrate the person back into the community. Prison involves a much more *coercive* effort, explicitly repressive in character, to segregate the person from the community. The development of probation as a cooptive strategy of social control requires further study: first, to determine how different its tactics⁵ were and are from the coercive use of imprisonment; second, to determine how and why probation emerged as a modern social control strategy; and third, to determine when and for whom this strategy was and is employed. Research directed to these issues should help to resolve further the question of what the effect of probation might conceivably have been in terms of imposing "industrial discipline." Alternatively, to assume all aspects of social control, be they cooptive or coercive, are equivalent, and to ignore the emergence of different types of social control under varying social conditions, is to neglect some of the most significant developments in contemporary responses to deviance. Our theories will benefit from an acknowledgment of change.

John Hagan
Jeffrey Leon
University of Toronto

REFERENCES

- Hagan, John
1977 *The Disreputable Pleasures*. Toronto: McGraw-Hill Ryerson.
- Hagan, John and Jeffrey Leon
1977a "Rediscovering delinquency: social history, political ideology and the sociology of law." *American Sociological Review* 42:587-98.
1977b "Philosophy and sociology of crime control: Canadian-American comparisons." *Sociological Inquiry* 47:181-208.
- Hagan, John, Edward Silva and John Simpson
1977 "Conflict and consensus and the designation of deviance." *Social Forces* 56:320-40.
- Leon, Jeffrey
1977 "The development of Canadian juvenile justice: a background for reform." *Osgoode Hall Law Journal* 15:71-106.

⁵ Note that in our research the tactics of probation include both official and unofficial activities of probation officers; that is, both formal and informal actions of formal control agents. Lauderdale and Larson's emphasis on the *occurrence* statistic, referring as it does to a wide range of unofficial contacts, underscores the need to make distinctions even among the cooptive measures.

- Platt, Anthony
 1969 *The Child Savers*. Chicago: University of Chicago Press.
 1973 "Dialogue with Anthony Platt." *Issues in Criminology* 8:19-33.
 1974 "The triumph of benevolence: the origins of the juvenile justice system in the United States." Pp. 356-89 in Richard Quinney (ed.), *Criminal Justice in America*. Boston: Little, Brown.
- Schultz, J. L.
 1973 "The cycle of juvenile court history." *Crime and Delinquency* 19:457-76.

ON THE USE OF HISTORICAL DOCUMENTS AND EVIDENCE

(COMMENT ON MELBIN, ASR FEBRUARY, 1978)*

A recent paper by Melbin (1978) explores the notion that nighttime social life in urban areas resembles social life on former land frontiers. While a fascinating idea, the empirical treatment of the central theme is seriously deficient. I would find little fault with the paper if it had been presented and elaborated only as a conceptual idea. However, the paper purports to *test* the hypothesis that nighttime urban social organization parallels frontier social organization with the presentation of *evidence*. It is with this presentation of evidence that I take issue in general and the use of historical materials in particular. Specifically, Melbin errs in the following ways: One, the documentation is stated to be archival in nature. This is not the case. Two, certain assumptions regarding presumed historical sequences are made which simply are not correct; nor is the use of sources with regard to those assumptions. The author exhibits a cavalier expediency with regard to historical data which both weakens a potentially useful argument as well as nullifies any test of relationships.

Several points should be considered. First, Melbin states that, "I have chosen to review life in the U. S. West in the middle of the nineteenth century along with the present-day nighttime. Of course there were other land frontiers and the hypothesis should apply to all of them" (Melbin, 1978:6).

A sufficient reading of the literature on frontier social structure would have demonstrated the falsity of this statement. For example the California gold rush camps (1848-1860) were

significantly different from the later Nevada silver camps, and the Colorado, Montana, and Yukon gold camps. The most obvious reason for this is the fact that in the latter rushes miners brought with them experience with regard to the technology of placer mining as well as camp life in general. The experiences of frontier farming communities also were significantly different than those of mining contexts. Frontier farming communities for example, had a much more equal distribution on the basis of sex and age. Mining communities tended to be almost exclusively young males. The former tended to be quite permanent while the latter were often "tent towns" which would literally vanish with the rumor of another gold (or silver) strike (Silver, 1978). The economic base of the two are different enough to require no elaboration. Therefore, although the data and experiences gathered from one source *may* be generalizable to all frontier contexts, the assumption cannot be that the data "should apply to all of them" (Melbin, 1978:6).

More importantly, Melbin next states, "However there are good reasons to begin by demonstrating it [the frontier hypothesis] for the U. S. West. One is that the archives holding information about this westward flow are thorough, well organized, and readily available" (Melbin, 1978:6). While this *may* be true in *some* instances (it is not generally the case) the number of sources that remotely could be construed as archival in this study number only five (out of 41 separate citations). These five are not original reports but are all found in edited composites. Seven other citations of that era are from an historian who, while possibly experiencing some of the patterns of which Melbin speaks, could not have been a firsthand source to all. The rest of the citations (from eight sources) were of contemporary writers—some of dubious authority.

The importance of accurate archival data to Melbin's argument cannot be overlooked. This "archival evidence" as presented, is no more than introductory sociology term paper tactics.

In the first discussion of similarities and differences between land frontiers and time frontiers—*Advance Is in Stages*—Melbin uses as historical evidence the following: "in a series of waves . . . the hunter and the fur trader who pushed into the Indian country were followed by the cattle raiser and he by the pioneer farmer" (Turner, cited by Melbin, 1978:6). And: "Life styles were distinctive in each stage as well. The hunters and trappers did not dwell like the miner who followed, and they in turn lived differently from the pioneer farmers who came later" (Billington, cited by Melbin, 1978:6).

* Address all communications to: Burton B. Silver; 1041 W. Camino Real; Boca Raton, FL 33432.

The point here is that farmers began settling the Oregon Territory (1830s–1840s) *before* gold was discovered in California (in 1848). Farmers also settled areas in the Midwest *before* the cattle raiser. We are in reality dealing with geographic variations rather than a firm sequence of stages. The sequence of settlement stages can only apply to extremely limited areas—and indeed limited aspects of time. Further explorations into the data sources stressed by Melbin surely would have uncovered this oversight (see Silver, 1978).

To illustrate further, in his discussion of violence (*There Is More Lawlessness and Violence*) the author states as evidence: "Yet the violence was concentrated in certain places; otherwise killings and mob law were remarkably infrequent. Such infamous towns as Tombstone and Deadwood, and the states of Texas and California had more than their share of gunfights" (Melbin, 1978:11). Melbin's "archival" sources are listed as: Frantz and Choate, 1955; Billington, 1949; Hollon, 1973—none of which are remotely archival. The reader is at a loss to know what is a fair share of gunfights. Melbin continues:

But the tumult in the cow towns was seasonal, and took place when the cowboys finally reached Abilene, Ellsworth, and Dodge City after the long drive. And the mayhem was selective. Flint (1826:401) wrote, "Instances of murder, numerous and horrible in their circumstances, have occurred in my vicinity . . . in which the drunkenness, brutality, and violence were mutual. . . [yet] quiet and sober men would be in no danger of being involved." (Melbin, 1978:11)

My point here is that there were no cow towns in 1826 as the source (Flint) would indicate. The towns, so called, became famous for the cattle drives only after the Civil War and when, in 1867, Abilene was established as a major shipping center. It was at this time that the cattle drives from Texas began along with the development of cow towns—made popular in films and television with distortions of the actual character of the places. Melbin's use of "firsthand" historical (or archival) sources does not correspond with the empirical point or time period at hand.

Another example of poor use of relevant time framework is found in his discussion, *New Behavioral Styles Emerge*. With reference to settlements remote from official courts, evidence emanates from the same above source: Flint, 1826. As quoted in Melbin (1978:10), Flint writes: "It is true that there are worthless people here. . . ." The "West," in 1826, was still essentially east of the Mississippi River. If Melbin is speaking of the West of the latter part of the nineteenth century (he never says, but

the reader must presume he is), then again the "evidence" presented is erroneous. Equally flagrant violations of evidence presentation are to be found in the other six discussions (pp. 7–19). Dwelling on these errors and oversights further seems unnecessary.

The use of historical materials cannot be taken lightly. Commendable use of historical materials has been made by Erikson (1966) and Kanter (1968) to name some of the more notable examples. The utility of this approach is to observe concepts and variables in a context of closure and control. The strength of an historical frame of reference is in our ability to step back and test our principles in a setting which has a definable beginning and a definable end. It is not unlike laboratory control. While there is inevitable error in these sources, the strength of the context which provides closure, control, and no recurrent input gives us yet another framework with which to understand the dynamics of society.

The empirical strength of Melbin's paper is in his field experiments (unfortunately, testing only one of ten propositions, however). There is, to be sure, no strength in his use of historical materials.

Direct data sources (diaries, journals, newspapers, census) are indeed all readily available. It is a shame that they were all overlooked. Some immediate sources which come to mind are the National Archives, Huntington Library, numerous historical societies of western states, microfilm copies of untold numbers of newspapers, etc., as well as *published, firsthand* diaries and journals of those who experienced the processes of which Melbin writes.

To reiterate, I am not in basic disagreement with the general thesis of the paper. My point is that at a time when sociologists are presumably still striving for empirical respectability, weak presentations appearing in the "flagship" journal should be avoided.

Burton B. Silver
Boca Raton, Florida

REFERENCES

- Erikson, Kai
1966 *Wayward Puritans*. New York: Wiley.
- Kanter, Rosabeth
1968 "Commitment and social organization; a study of commitment mechanisms in utopian communities." *American Sociological Review* 33:499–516.
- Melbin, Murray
1978 "Night as frontier." *American Sociological Review*: 43:3–22.

Silver, Burton

1978 "Social structure and legal evolution: a longitudinal analysis of the California gold rush." Paper presented at the annual meeting of the American Sociological Association, San Francisco.

REPLY TO SILVER: ON THE STRUCTURE OF EVIDENCE*

Although he does not disagree with my hypothesis of "Night as Frontier," my critic claims that (1) frontiers are more varied and not as uniform as I described them, (2) I did not use as much archival evidence as I said I would (I did not say so), and (3) my more contemporary sources either are used wrongly or are of dubious authority.

My portrait of the U.S. western frontier covered a fifty-year period, "Specifically . . . the middle decades of the nineteenth century, about 1830-1880" (1978:6). I chose this wide span to avoid dwelling on an outstanding feature that was more temporary, such as the fur trade of the 1830s in the northwest territories, the California gold rush of 1848-1852, or building the Union Pacific railroad in the 1860s. Understandably there are variations throughout five decades and in different places. In making the general case I discarded some differences, but my ten propositions still hold. For example, even farming communities were *initially* populated by young adult males. Men often left their families behind and came west to "spy out the land," and having made a new home they then sent for their wives and children (Dick, 1937:7,18). Concerning the sequence of stages: where farmers settled before the cattle raiser the cattle raising stage was *skipped*. It is not that the reverse occurred—that farmers moved out and then cattlemen moved in. And in my paragraph saying that violence was concentrated in certain places (1978:11), of the five items of testimony covering the fifty-year span only two pertain to cow towns. Elsewhere (1978:9) I did err chronologically in offering a quotation from William Penn, who lived much earlier.

Riegel (1947:1) said that "Some day there may be a study of frontier advance which will show the traits common to all frontiers." In that spirit I thought it more important to reconstruct the main social patterns of a half-century than to blur them with notes on variations. I make an equally broad claim about expansion

of wakeful activity in the nighttime. It applies to Singapore, Europe, Peru, and the U.S., although I know there will be local variations among these too.

About the adequacy of my tests, and especially of the historical evidence offered: the more general a hypothesis, the greater the variety of potential disconfirmation or support for it. There is more information to scan, more diverse tests to apply, more chances for battering the hypothesis with inconsistencies among the findings. Consistent support from varied sources would convey validity strongly for the idea. So I collected many kinds of data and made my case depend on collateral information corroborating each other. I used analyses by historians (propositions 1 through 10), census data (proposition 2), sampled observations (propositions 2,8), archival materials such as letters and autobiographies by travelers and frontiersmen (i.e., participant observations, propositions 3,6,7,8), an unobtrusive measure (reports of fights in emergency calls to telephone operators, proposition 7), experiments (proposition 8), and newspaper reports (proposition 10).

Why the fuss about archives? Partly it comes from misunderstanding a comment in the essay. I did not say that my documentation would be "archival in nature." I said that one of the reasons for studying the U.S. West was that the archives holding information about it were thorough, well-organized, and readily available, compared with what was available to me about other frontiers (such as Russia east of the Urals, or the Australian Outback) (1978:6,20).

Partly it comes from an unwillingness to accept firsthand accounts if they are reprinted in "edited composites." That is like insisting a student of Shakespeare should work only with a first folio edition and not with duplicates of it. I welcomed anthologies containing earlier-published autobiographical materials. The firsthandness of those accounts were not diminished by *reprinting*. Moreover, a number of times I summarized historians' works that were themselves based on references to archives. Here are instances: Hollon's analysis of frontier violence (my p. 11) used Robert Dykstra's *The Cattle Towns*. Dykstra compiled his statistics by studying the local newspapers of the era. The quotation about the Homestead Act (my p.17) was drawn by Smith from the first session of the thirty-sixth congress, *Congressional Globe* (April 10, 1860). I provided page references so that a reader interested in more details could promptly follow up that interest.

Partly, it comes from rejecting prior historical

* Address all communications to: Murray Melbin; College of Liberal Arts; Boston University; 96-100 Cummington St.; Boston, MA 02215.

scholarship as a source of evidence. I sought to characterize a land frontier, and to summarize its features fairly and generally enough so that I could reconceptualize the idea to fit a space-time container of living rather than only a spatial one. Hence I read much about the U.S. West, far more than the score of writings cited in my essay. From that review I judged that historians such as Turner, Dick, Riegel, Billington, Smith and Hollon provided sound information. (There is controversy over some of Turner's notions about the American frontier but not involving my ten propositions.) Researchers in every field rely on the prior organized work of others. We use scholarly publications to save the effort of always returning to beginnings. And it fits the principle of science as social, as an ongoing cooperative venture for cumulative knowledge. The conclusions of twentieth-century historians are in some ways more authoritative than their predecessors (or than particular archives) because our contemporaries are helped by the efforts of earlier scholars along with directly marshalling and sifting through more archives. Accordingly, in the "Night as Frontier" essay the

structure of evidence is a systematic integration of my theoretical reasoning and new data with the achievements of historians.

The names of "contemporary writers of dubious authority" are not mentioned so I cannot assess that claim. But assuming that my critic has fulfilled his research tasks well in studying western mining camps I look forward to learning more about gold from Silver.

Murray Melbin
Boston University

REFERENCES

- Dick, Everett
[1937] *The Sod-House Frontier 1854-1890*. New York: Appleton-Century.
Dykstra, Robert
1968 *The Cattle Towns*. New York: Knopf.
Melbin, Murray
1978 "Night as frontier." *American Sociological Review* 43:3-22.
Riegel, Robert E.
1947 *America Moves West*. New York: Holt.

ITEMS (Continued)

Associate Editors: Bernard Farber, Daniel Glazer, Gisela Hinkle, James R. Kluegel, John Lofland, Valerie Oppenheimer, Alejandro Portes, Barbara Reskin, Richard Rubinson, John Wilson, Mayer Zald.

■ **AMOS H. HAWLEY** (Cumulative Change in Theory and History) is Kenan Professor Emeritus in the Department of Sociology at the University of North Carolina. His research focuses on multinational fertility and migration. He also is studying migration and employment status in Malaysia. In 1975 he coedited (with Vincent Rock) *Metropolitan America in Contemporary Perspective* (Wiley).

■ **JOHN BOLI-BENNETT** (Ideology of Childhood and the State) is Visiting Assistant Professor in the Department of Sociology and Anthropology at the University of Santa Clara. His research interests include an historical study of the formation and propagation of dominant world-system ideologies, especially socialism; a general review of national planning in the twentieth century; and a study of the power and autonomy of the state in national societies. **JOHN W. MEYER** is Professor in the Department of Sociology and the Boys' Town Center for Youth Development at Stanford University. He is doing comparative studies of structural change in contemporary states and societies and studies of formal organizations as internally-decoupled reflections of institutional structures. He has coauthored (with Michael Hannan) *National Development in the World System* (University of Chicago Press, 1979).

■ **ROSS M. STOLZENBERG** (Bringing the Boss Back In) is Associate Professor of Sociology in the Department of Sociology, University of Illinois at Urbana-Champaign. His research centers on employer and industry effects on earnings, occupational attainment, working conditions, and on race and gender inequality in the labor market.

■ **JAMES R. LINCOLN** (Cultural Effects on Organizational Structure) is Associate Professor in the Department of Sociology at Indiana University. He continues to study (with Jon Olson and Mitsuyo Hanada) Japanese firms in the United States. He also is investigating interorganizational networks of juvenile justice system agencies. **JON OLSON** is a Ph.D. candidate in Sociology at the University of Southern California. Besides his work with Lincoln and Hanada, he is studying the organizational bases of social stratification. **MITSUYO HANADA** is Associate Professor in the Institute of Business Administration and Management (Sangyo Nohritsu Daigaku), Tokyo, Japan. Hanada is conducting a study of employees' commitment to organization in Japanese business subsidiaries in Thailand.

■ **MICHEAL W. GILES** (White Enrollment Stability) is Associate Professor in the Department of Political Science at Florida Atlantic University. Currently he is examining the effects of contextual/structural variables on the processes of group identification and individual-level racial attitudes. His book *School Desegregation Policy: Resistance, Acceptance, and the Alternatives* (coauthored with Douglas S. Gatlin and Everett F. Cataldo, D. C. Heath, forthcoming), summarized his study of individual-level compliance with school desegregation.

■ **PAUL D. ALLISON** (Measures of Inequality) is Assistant Professor in the Department of Sociology at Cornell University. His research interests include the interrelationships among sponsorship, collaboration, institutional prestige and productivity in the careers of scientists, and stochastic process models of measurement error.

■ **STEVEN STACK** (Direct Government Involvement in the Economy) is Assistant Professor of Sociology at Alma College. His research projects include studies of the impact of economic factors on the rate of suicide, the effect of clearance rates on homicide rates, and a test of a Keynesian model of income inequality.

■ **J. SCOTT LONG** (Productivity and Academic Position) is Assistant Professor of Sociology at Washington State University. His research focuses on career processes in scientific careers and on the effects of institutional location on scientific productivity.

■ **STEVEN D. McLAUGHLIN** (Occupational Sex Identification) is Assistant Professor of Sociology at Utah State University. He is investigating the conditions under which female labor force participation is responsive to changes in family composition.

■ **PAT LAUDERDALE** (Comment on Hagan and Leon, ASR August, 1977) is Assistant Professor of Sociology at the University of Minnesota. His research includes a comparative, longitudinal analysis of changing definitions of deviance and the role of political entities in boundary maintenance/movement. **GERALD LARSON** is a Ph. D. candidate in the Department of Sociology, University of Minnesota. He is interested in social solidarity and deviance.

■ **BURTON B. SILVER** (Comment on Melbin, ASR February, 1978) is analysing the evolution of legal systems. He also is studying the ephemeral role in contemporary communication systems.

1978
EDITORIAL CONSULTANTS

Akers, Ronald L.	Cannell, Charles	Erlanger, Howard S.
Albrecht, Gary L.	Caplow, Theodore	Evans, Robert R.
Aldous, Joan	Carey, Alex	Evers, Mark
Aldrich, Howard E.	Catton, William R.	Fararc, Thomas J.
Alexander, C. Norman, Jr.	Chambliss, William J.	Farber, Bernard
Alexander, Karl L.	Chase, Allan	Farley, W. Reynolds
Alford, Robert R.	Chase, Ivan	Faunce, William A.
Allen, Michael P.	Chase-Dunn, Christopher	Feagin, Joe R.
Alwin, Duane F.	Cheatwood, A. Derral	Featherman, David L.
Anderson, Barbara A.	Cherlin, Andrew J.	Fendrich, James M.
Anderson, James G.	Choldin, Harvey M.	Fennessey, James J.
Armer, J. Michael	Cicourel, Aaron V.	Ferree, Myra Marx
Arnold, David O.	Clark, John P.	Fine, Gary A.
Aveni, Adrian	Clemente, Frank	Fischer, Claude
Babchuk, Nicholas	Clogg Clifford	Fishman, Joshua
Bach, Kurt	Cloward, Richard A.	Fligstein, Neil
Bailey, Kenneth D.	Cockerham, William C.	Form, William H.
Baker, Paul J.	Cohen, Lawrence E.	Fox, John D.
Baldassare, Mark	Cohen, Roberta	Franks, David D.
Baldwin, John D.	Cole, Jonathan R.	Freeman, Howard E.
Barchas, Patricia R.	Cole, Stephen	Freidson, Eliot
Bean, Frank D.	Coleman, James S.	Frisbie, Parker
Becker, Howard S.	Collins, Randall	Fuguitt, Glenn V.
Bell, Wendell	Conner, Walter A.	Gale, Richard P.
Bengston, Vern	Converse, Philip E.	Gamson, William A.
Bennett, H. Stith	Cook, Karen	Gans, Herbert J.
Benson, J. Kenneth	Coser, Lewis	Garner, Roberta Ash
Berger, Bennett M.	Costner, Herbert L.	Gelles, Richard J.
Berk, Richard A.	Couch, Carl J.	Germani, Gino
Bernard, Jessie	Crain, Robert	Gibbs, Jack P.
Biddle, Bruce J.	Cramer, M. Richard	Glenn, Norval D.
Bielby, William T.	Crane, Diana M.	Glock, Charles Y.
Birnbaum, Norman	Cutright, Phillips	Goldstein, Sidney
Bittner, Egon	D'Antonio, William V.	Goode, William J.
Blau, Peter M.	D'Amico, Ronald J.	Gordon, Robert A.
Blumer, Herbert	Darrow, William	Gove, Walter R.
Boguslaw, Robert	Davis, Fred	Granovetter, Mark
Bohrnstedt, George	Davis, James A.	Grasmick, Harold
Bonacich, Edna	Davis, Kingsley	Greeley, Andrew M.
Bonjean, Charles M.	Davis, Murray	Gregory, Stanford W.
Bordua, David J.	Davis, Nanette J.	Griffin, Larry J.
Bose, Christine E.	Delacroix, Jacques	Grimshaw, Allen D.
Boyd, John	Denzin, Norman K.	Guest, Avery M.
Brede, Richard	Dibble, Vernon	Gusfield, Joseph R.
Breiger, Ronald L.	Dillingham, Gerald	Gutman, Robert
Brymer, Richard	Dinitz, Simon	Hadden, Jeffrey K.
Buckley, Walter	Drabek, Thomas E.	Hage, Jerald T.
Bumpass, Larry L.	Dubin, Robert	Hagstrom, Warren O.
Burger, Thomas	Dutton, Diana	Halaby, Charles N.
Burgess, Robert	Eckland, Bruce K.	Hall, Peter M.
Burke, Peter J.	Edmonston, Barry	Hall, Richard H.
Burstein, Paul	Eitzen, D. Stanley	Haller, Archibald O.
Burt, Ronald S.	Elder, Glen Holl, Jr.	Hallinan, Maureen T.
Cain, Pamela S.	Emerson, Richard M.	Hamblin, Robert L.
Camilleri, Santo F.	Empey, LaMar T.	Hannan, Michael T.
Campbell, Richard T.	Erickson, Kai T.	Hargens, Lowell L.

- Hauser, Robert M.
 Hawkins, Richard
 Hazelrigg, Lawrence
 Heer, David M.
 Heirich, Max A.
 Heise, David
 Hendershot, Gerry
 Henshel, Richard L.
 Hepburn, John R.
 Hewitt, Christopher
 Heyl, Barbara
 Hill, Richard
 Hinkle, Gisela
 Hinkle, Roscoe C.
 Hirsch, Jerry
 Hirsch, Paul M.
 Hirschi, Travis W.
 Hochschild, Arlie R.
 Horowitz, Irving Louis
 Horwiz, Allan V.
 House, James S.
 Hout, Michael
 Huber, Joan
 Hudis, Paula M.
 Hummon, Norman P.
 Hymes, Dell
 Jackman, Mary R.
 Jackman, Robert W.
 Jencks, Christopher
 Jenkins, J. Craig
 Jesser, Clinton J.
 Johnson, Barclay
 Johnson, Benton
 Johnson, Michael P.
 Johnson, Norris R.
 Jones, Robert A.
 Kadushin, Charles
 Kalleberg, Arne L.
 Karsh, Bernard
 Kasarda, John D.
 Kennedy, Robert E., Jr.
 Kervin, John B.
 Killian, Lewis M.
 Kimberly, John
 King, David
 Klapp, Orrin
 Klein, Malcolm W.
 Kluegel, James R.
 Knoke, David
 Kohn, Melvin
 Kohout, Frank J.
 Komarovsky, Mirra
 Korpi, Walter
 Kraus, Richard
 Ladinsky, Jack
 Ladner, Joyce
 Land, Kenneth
 Landsberger, Henry A.
 Lantz, Herman
 Laumann, Edward O.
 Leik, Robert
 Leinhardt, Sam
 Lenski, Gerhard, Jr.
 Levine, Joel H.
 Lewis, J. David
 Lieberman, Stanley
 Lipset, Seymour Martin
 Liska, Allen E.
 Litwak, Eugene
 Lofland, John
 Long, Larry
 McAuliffe, William
 McCall, George J.
 McCarthy, John
 McClelland, David
 McClendon, McKee
 McFarland, David D.
 McPhail, Thomas
 Maines, David R.
 Mare, Robert D.
 Maris, Ronald W.
 Marks, Stephen
 Marsden, Peter V.
 Marwell, Gerald
 Marx, Gary T.
 Mason, Karen Oppenheim
 Mauss, Armand L.
 Maxwell, William
 Mayhew, Bruce H.
 Mazur, Allan
 Meeker, Barbara F.
 Meier, Robert
 Meltzer, Bernard N.
 Mencken, Jane
 Mercer, Jane R.
 Merton, Robert K.
 Meyer, John W.
 Meyer, Marshall W.
 Michelson, William
 Micklin, Michael
 Miller, David
 Miller, Jerry L.L.
 Mills, Theodore
 Moch, Michael
 Modigliani, Andre
 Molotch, Harvey L.
 Moore, Kristin
 Mullins, Nicholas C.
 Myers, George C.
 Nahemow, Nina R.
 Nettler, Gwynn
 Oppenheimer, Valerie K.
 Ornstein, Michael
 Orum, Anthony M.
 O'Toole, Richard
 Pampel, Fred C.
 Parish, William L.
 Parsons, Talcott
 Payne, David E.
 Pearlman, Leonard
 Perinbanayagam, R. S.
 Perrin, Robert G.
 Perrow, Charles
 Peterson, Richard A.
 Petras, John W.
 Pettigrew, Thomas F.
 Pfeffer, Jeffrey
 Pine, Vanderlyn
 Ploch, Donald R.
 Pondy, Louis R.
 Portes, Alejandro
 Psathas, George
 Quadagno, Jill
 Ragin, Charles
 Rapoport, Anatol
 Rau, William
 Reeder, Leo G.
 Reiss, Albert J., Jr.
 Reiss, Ira L.
 Reskin, Barbara
 Ridgeway, Cecelia
 Riecken, Henry
 Riley, Matilda
 Rindfuss, Ronald R.
 Ritzer, George
 Robinson, John P.
 Robinson, Robert
 Robinson, Sherman
 Rokeach, Milton
 Roof, W. Clark
 Rose, Ed
 Rosenfeld, Rachel A.
 Rossi, Alice S.
 Rossi, Peter H.
 Roth, Guenther
 Roth, Julius A.
 Rubinson, Richard B.
 Rushing, William A.
 Ryder, Norman B.
 Scheff, Thomas J.
 Schegloff, Emanuel A.
 Schneider, Louis
 Schoen, Robert
 Schoenberg, Ronald J.
 Schooler, Carmi
 Schuessler, Karl F.
 Schulman, Gary
 Schur, Edwin M.
 Schwartz, Richard
 Scott, W. Richard
 Seeman, Melvin
 Sennett, Richard
 Sheingold, Carl
 Shephard, Jon M.
 Short, James F.
 Shotland, Lance
 Siegel, Paul M.
 Simkus, Albert
 Simon, William
 Simpson, Richard L.

- Singer, Benjamin D.
Skolnick, Arlene
Skolnick, Jerome H.
Smith, David H.
Smith, Don D.
Smucker, Joseph
Snyder, Charles R.
Southwood, Kenneth
Spaeth, Joe L.
Speare, Alden, Jr.
Spector, Malcolm
Spilerman, Seymour
Stack, Steven
Stark, Rodney W.
Stewart, Robert L.
Stewman, Shelby
Stryker, Sheldon
Sudman, Seymour
Swafford, Michael S.
Sweet, James A.
Szymanski, Albert J.
Taeuber, Karl E.
Terry, Robert M.
Tessler, Richard
Thorne, Barrie
Tiryakian, Edward
- Tittle, Charles R.
Tomasson, Richard F.
Treas, Judith
Tucker, Charles W.
Turk, Austin T.
Turner, Jonathan
Turner, Ralph H.
Turner, R. Jay
Tyree, Andrea
Useem, Michael
Van den Berghe, Pierre
Vander Kooi, Ronald
van de Walle, Etienne
Vanreman, Reeve
Vernon, Glenn M.
Villemez, Wayne J.
Vogel, Ezra
Voss, Paul
Wahrman, Ralph
Waite, Linda J.
Wallerstein, Immanuel
Walton, John
Wanderer, Jules
Warren, Donald
Webster, Murray A., Jr.
Weiland, Marion W.
- Weller, Jack M.
Westoff, Charles F.
White, Harrison C.
Whyte, Martin King
Wieder, Lawrence
Wiggins, James A.
Wiley, Norbert F.
Willer, David E.
Williams, Robin M., Jr.
Wilson, Everett K.
Wilson, Franklin D.
Wilson, James Q.
Wilsor, John
Wilsor, William J.
Wiseman, Jacqueline P.
Wolf, Wendy
Wright, Erik O.
Wright, James D.
Yancey, William L.
Yinger, J. Milton
Zald, Mayer N.
Zeitlin, Maurice
Zimmer, Basil G.
Zucker, Lynne G.
Zuckerman, Harriet
Zurcher, Louis A.

INDEX

AMERICAN SOCIOLOGICAL REVIEW

Published Bimonthly by the American Sociological Association

CONTENTS OF VOLUME 43, NUMBERS 1-6

ARTICLES

Adler, Israel and Judith A. Shuval Cross Pressures during Socialization for Medicine	693
Alexander, Jeffrey C. Formal and Substantive Voluntarism in the Work of Talcott Parsons: A Theoretical and Ideological Reinterpretation	177
Alexander, Karl L., Martha Cook, and Edward L. McDill Curriculum Tracking and Educational Stratification: Some Further Evidence	47
Allison, Paul D. Measures of Inequality	865
Appelbaum, Richard P. Marx's Theory of the Falling Rate of Profit: Towards a Dialectical Analysis of Structural Social Change	67
Armer, Michael, and Larry Isaac Determinants and Behavioral Consequences of Psychological Modernity: Empirical Evidence from Costa Rica	316
Aronoff, Joel See William D. Crano	463
Baldwin, Janice I. See John D. Baldwin	335
Baldwin, John D., and Janice I. Baldwin Behaviorism on Verstehen and Erklären	335
Baumgartner, Robert See Thomas A. Heberlein	447
Beck, E.M., Patrick M. Horan, and Charles M. Tolbert II Stratification in a Dual Economy: a Sectoral Model of Earnings Determination	704
Bell, Wendell See Robert V. Robinson	125
Boli-Bennett, John and John Meyer The Ideology of Childhood and the State: Rules Distinguishing Children in National Constitutions, 1870-1970	797
Cohen, Lawrence E., and James R. Kluegel Determinants of Juvenile Court Dispositions: Ascriptive and Achieved Factors in Two Metropolitan Courts	162
Cook, Karen, and Richard M. Emerson Power, Equity and Commitment in Exchange Networks	721
Cook, Martha See Karl L. Alexander	47
Crano, William D., and Joel Aronoff A Cross-Cultural Study of Expressive and Instrumental Role Complementarity in the Family	463
Danigelis, Nicholas L. Black Political Participation in the United States: Some Recent Evidence	756
Driskell, James E., Jr. See Murray Webster, Jr.	220
Dutton, Diana B. Explaining the Low Use of Health Services by the Poor: Costs, Attitudes, or Delivery Systems?	348
Eaton, William W., and Andrée Fortin A Third Interpretation for the Generating Process of the Negative Binomial Distribution	264
Eder, Donna, and Maureen T. Hallinan Sex Differences in Children's Friendships	237
Emerson, Richard M. See Karen Cook	721
Firebaugh, Glenn A Rule for Inferring Individual-Level Relationships from Aggregate Data	557
Fortin, Andrée See William W. Eaton	264
Franke, Richard H., and James D. Kaul The Hawthorne Experiments: First Statistical Interpretation	623
Friedland, Roger See Alexander M. Hicks	302
Gartrell, John W. See R. Jay Turner	368
Giles, Micheal W. White Enrollment Stability and School Desegregation: A Two Level Analysis	848
Guest, Avery M. Suburban Social Status: Persistence or Evolution?	251
Hallinan, Maureen T. See Donna Eder	237
Hamblin, Robert L. See Brian L. Pitcher	23
Hanada, Mitsuyo See James R. Lincoln	829
Hawley, Amos Presidential Address: Cumulative Change in Theory and in History	787
Heberlein, Thomas A., and Robert Baumgartner Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature	447

Hicks, Alexander, Roger Friedland, and Edwin D. Jonnson Class Power and State Policy: The Case of Large Business Corporations, Labor Unions and Governmental Redistribution in the American States	302
Hindelang, Michael Race and Involvement in Common Law Personal Crimes	93
Hogan, Dennis P. The Variable Order of Events in the Life Course	573
Horan, Patrick M. See E. M. Beck	534
Horan, Patrick M. Is Status Attainment Research Atheoretical?	704
Isaac, Larry See Michael Armer	316
Johnson, Edwin See Alexander M. Hicks	302
Kaul, James D. See Richard H. Franke	623
Kluegel, James R. See Lawrence E. Cohen	162
Kluegel, James R. The Causes and Cost of Racial Exclusion from Job Authority ..	285
Lever, Janet Sex Differences in the Complexity of Children's Play and Games	471
Lincoln, James R. Community Structure and Industrial Conflict: An Analysis of Strike Activity in SMSAs	199
Lincoln, James R., Jon Olson and Mitsuyo Hanada Cultural Effects on Organizational Structure: The Case of Japanese Firms in the U.S.	829
Long, J. Scott Productivity and Academic Position in the Scientific Career	889
McDill, Edward L. See Karl L. Alexander	47
McLaughlin, Steven D. Occupational Sex Identification and the Assessment of Male and Female Earnings Inequality	909
Marini, Margaret Mooney The Transition to Adulthood: Sex Differences in Educational Attainment and Age at Marriage	483
Melbin, Murray Night As Frontier	3
Meyer, John See John Boli-Bennett	797
Miller, Jerry L. L. See Brian L. Pitcher	23
Myles, John F. Institutionalization and Sick Role Identification among the Elderly .	508
Olson, Jon See James R. Lincoln	829
Payne, David E. Cross-National Diffusion: The Effects of Canadian TV on Rural Minnesota Viewers	740
Pitcher, Brian L., Robert L. Hamblin, and Jerry L. L. Miller The Diffusion of Collective Violence	23
Reed, Theodore L. Organizational Change in the American Foreign Service, 1925-1965: The Utility of Cohort Analysis	404
Robinson, Robert V., and Wendell Bell Equality, Success, and Social Justice in England and the United States	125
Rosenblum, Barbara Style As Social Process	422
Rosenfeld, Rachel A. Women's Intergenerational Occupational Mobility	36
Rushing, William A. Status Resources, Societal Reaction, and Type of Mental Hospital Admission	521
Scarr, Sandra, and Richard A. Weinberg The Effects of "Family Background" on Intellectual Attainment	674
Shuval, Judith A. See Israel Adler	693
Simkus, Albert A. Residential Segregation by Occupation and Race in Ten Urbanized Areas, 1950-1970	81
Singer, Eleanor Informed Consent: Consequences for Response Rate and Response Quality in Social Surveys	144
Smith-Lovin, Lynn and Ann R. Tickamyer Nonrecursive Models of Labor Force Participation, Fertility Behavior and Sex Role Attitudes	541
Smith, Douglas A. See Charles R. Tittle	643
Stack, Steven The Effect of Direct Government Involvement in the Economy on the Degree of Income Inequality: A Cross-National Study	880
Stolzenberg, Ross M. Bringing the Boss Back In: Employer Size, Employee Schooling, and Socioeconomic Achievement	813
Swafford, Michael Sex Differences in Soviet Earnings	657
Tickamyer, Ann R. See Lynn Smith-Lovin	541
Tittle, Charles R., Wayne J. Villemez, and Douglas A. Smith The Myth of Social Class and Criminality: An Empirical Assessment of the Empirical Evidence	643
Tolbert, Charles M., II See E. M. Beck	704
Turner, R. Jay, and John W. Gartrell Social Factors in Psychiatric Outcome: Toward the Resolution of Interpretive Controversies	368


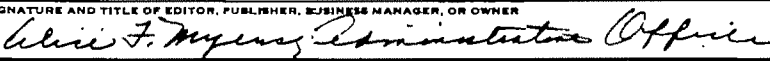
INDEX

939

Villemez, Wayne J. See Charles R. Tittle	643
Webster, Murray, Jr., and James E. Driskell, Jr. Status Generalization: A Review and Some New Data	220
Weinberg, Richard A. See Sandra Scarr	674
Wheaton, R. Blair The Sociogenesis of Psychological Disorder: Reexamining the Causal Issues with Longitudinal Data	383

COMMENTS

Burgess, Robert L. Reply to Pierce and Harrell	597
Butler, John S. Military Inequality: Reply to Hauser	607
Chase-Dunn, Christopher Reply to Stumpp, Marsh and Lake	604
Cohen, Jack See Charles F. Cortese	590
Cortese, Charles F., R. Frank Falk and Jack Cohen Understanding the Standardized Index of Dissimilarity: Reply to Massey	590
Falk, R. Frank See Charles F. Cortese	590
Grasmick, Harold G. and Steven D. McLaughlin Deterrence and Social Control: Comment on Silberman	272
Hagan, John and Jeffrey Leon Crass Conflict: A Reply to Lauderdale and Larson	925
Harrell, W. Andrew See W. David Pierce	597
Harry, Joseph and Mary C. Sengstock Attribution Goals and Deviance: Comment on Harris	278
Hauser, Robert M. On "Inequality in the Military" Comment on Butler	115
Hindelang, Michael See Travis Hirschi	610
Hirschi, Travis and Michael Hindelang Reply to Ronald L. Simons	610
Hope, Keith Diamonds are Trumps: Reply to House	615
House, James S. Some Facets and Flaws of the Hope Diamond: Comment on Hope	439
Jenkins, J. Craig Reply to Walsh	605
Lake, Deborah See Mark Stumpp	600
Larson, Gerald See Pat Lauderdale	922
Lauderdale, Pat and Gerald Larson Marxist and "Organizational" Approaches to Delinquency and the Sociology of Law: Crucial Problems in "Testing" the Perspectives: Comment on Hagan and Leon	922
Leon, Jeffrey See John Hagan	925
Levine, Joel H. Comparing Models of Mobility: Comment on McCann	118
Lincoln, James R. and Jon Olson The Ecological Approach and Community Leadership: Comment on Grimes, Bonjean, Lyon, and Lineberry	112
McLaughlin, Steven D. See Harold G. Grasmick	272
Marsh, Robert See Mark Stumpp	600
Massey, Douglas S. On the Measurement of Segregation as a Random Variable: Comment on Cortese, Falk and Cohen	587
Melbin, Murray Reply to Silver: On the Structure of Evidence	931
Olson, Jon See James Lincoln	112
Pierce, W. David and W. Andrew Harrell Threats to Internal Validity in Burgess and Nielsen's Social Exchange Experiment Comment on Burgess and Nielsen	572
Rubinson, Richard Reply to Stack	614
Sengstock, Mary C. See Joseph Harry	278
Silver, Burton On the Use of Historical Documents and Evidence: Comment on Melbin	929
Simons, Ronald L. The Meaning of IQ Delinquency: Comment on Hirschi and Hindelang	268
Stack, Steven Internal Political Organization and the World Economy of Income Inequality: Comment on Rubinson	271
Stumpp, Mark, Robert M. Marsh and Deborah A. Lake The Effects of International Economic Dependence on Development: A Critique: Comment on Chase-Dunn	600
Szymanski, Albert White Workers' Loss from Racial Discrimination: Reply to Villemez	776
Villemez, Wayne J. Black Subordination and White Economic Well-Being	772
Walsh, Edward J. On the Interaction between a Movement and Its Environment: Comment on Jenkins and Perrow	110

U.S. POSTAL SERVICE STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION <small>(Required by 39 U.S.C. 3685)</small>		
1. TITLE OF PUBLICATION American Sociological Review		A. PUBLICATION NO. 8 3 3 8 0 0
2. DATE OF FILING 9/30/78		
3. FREQUENCY OF ISSUE Bi-Monthly	A. NO. OF ISSUES PUBLISHED ANNUALLY 6	B. ANNUAL SUBSCRIPTION PRICE \$30.00
4. LOCATION OF KNOWN OFFICE OF PUBLICATION (Street, City, County, State and ZIP Code) (Not printers) 1722 N St., NW, Washington, DC 20036		
5. LOCATION OF THE HEADQUARTERS OR GENERAL BUSINESS OFFICES OF THE PUBLISHERS (Not printers) 1722 N St., NW, Washington, DC 20036		
6. NAMES AND COMPLETE ADDRESSES OF PUBLISHER, EDITOR, AND MANAGING EDITOR		
PUBLISHER (Name and Address) American Sociological Ass'n., 1722 N St., NW, Washington, DC 20036		
EDITOR (Name and Address) Rita J. Simon, Dept. of Sociology, Univ. of Illinois, Urbana, IL 61801		
MANAGING EDITOR (Name and Address)		
7. OWNER (If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of the individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given.)		
NAME	ADDRESS	
American Sociological Association	1722 N Street, NW	
	Washington, DC 20036	
8. KNOWN BONDHOLDERS, MORTGAGEES, AND OTHER SECURITY HOLDERS OWNING OR HOLDING 1 PERCENT OR MORE OF TOTAL AMOUNT OF BONDS, MORTGAGES OR OTHER SECURITIES (If there are none, so state)		
NAME	ADDRESS	
None		
9. FOR COMPLETION BY NONPROFIT ORGANIZATIONS AUTHORIZED TO MAIL AT SPECIAL RATES (Section 132.123, FPM) The purpose, function, and nonprofit status of this organization and the exempt status for Federal income tax purposes (Check one)		
<input checked="" type="checkbox"/> HAVE NOT CHANGED DURING PRECEDING 12 MONTHS <input type="checkbox"/> HAVE CHANGED DURING PRECEDING 12 MONTHS (If changed, publisher must submit explanation of change with this statement.)		
10. EXTENT AND NATURE OF CIRCULATION	AVERAGE NO. COPIES EACH ISSUE DURING PRECEDING 12 MONTHS	ACTUAL NO. COPIES OF SINGLE ISSUE PUBLISHED NEAREST TO FILING DATE
A. TOTAL NO. COPIES PRINTED (Net Press Run)	17,050	17,050
B. PAID CIRCULATION 1. SALES THROUGH DEALERS AND CARRIERS, STREET VENDORS AND COUNTER SALES	None	None
2. MAIL SUBSCRIPTIONS	14,852	15,402
C. TOTAL PAID CIRCULATION (Sum of 10B1 and 10B2)	14,852	15,402
D. FREE DISTRIBUTION BY MAIL, CARRIER OR OTHER MEANS SAMPLES, COMPLIMENTARY, AND OTHER FREE COPIES	None	None
E. TOTAL DISTRIBUTION (Sum of C and D)	14,852	15,402
F. COPIES NOT DISTRIBUTED 1. OFFICE USE, LEFT OVER, UNACCOUNTED, SPOILED AFTER PRINTING	2,198	1,648
2. RETURNS FROM NEWS AGENTS	None	None
G. TOTAL (Sum of E, F1 and F2—should equal net press run shown in A)	17,050	17,050
11. I certify that the statements made by me above are correct and complete.	SIGNATURE AND TITLE OF EDITOR, PUBLISHER, BUSINESS MANAGER, OR OWNER  Alice F. Myers Administrative Officer	
12. FOR COMPLETION BY PUBLISHERS MAILING AT THE REGULAR RATES (Section 132.121, Postal Service Manual)		
39 U. S. C. 3626 provides in pertinent part: "No person who would have been entitled to mail matter under former section 4359 of this title shall mail such matter at the rates provided under this subsection unless he files annually with the Postal Service a written request for permission to mail matter at such rates." In accordance with the provisions of this statute, I hereby request permission to mail the publication named in item 1 at the phased postage rates presently authorized by 39 U. S. C. 3626.		
SIGNATURE AND TITLE OF EDITOR, PUBLISHER, BUSINESS MANAGER, OR OWNER		
 Alice F. Myers Administrative Officer		